GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: February 8, 2002, 10:53:11 ; Search time 19.49 Seconds
(without alignments)
130.962 Million cell updates/sec

Title: US-10-006-430-51
Perfect score: 29
Sequence: 1 ccgaatgccacgaacgaacgaacgaacgaa 29

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472130 seqs, 824859755 residues

Total number of hits satisfying chosen parameters: 5886436

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_or:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_mus:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_mus:*
35: em_htg_rod:*
36: em_htg_other:*

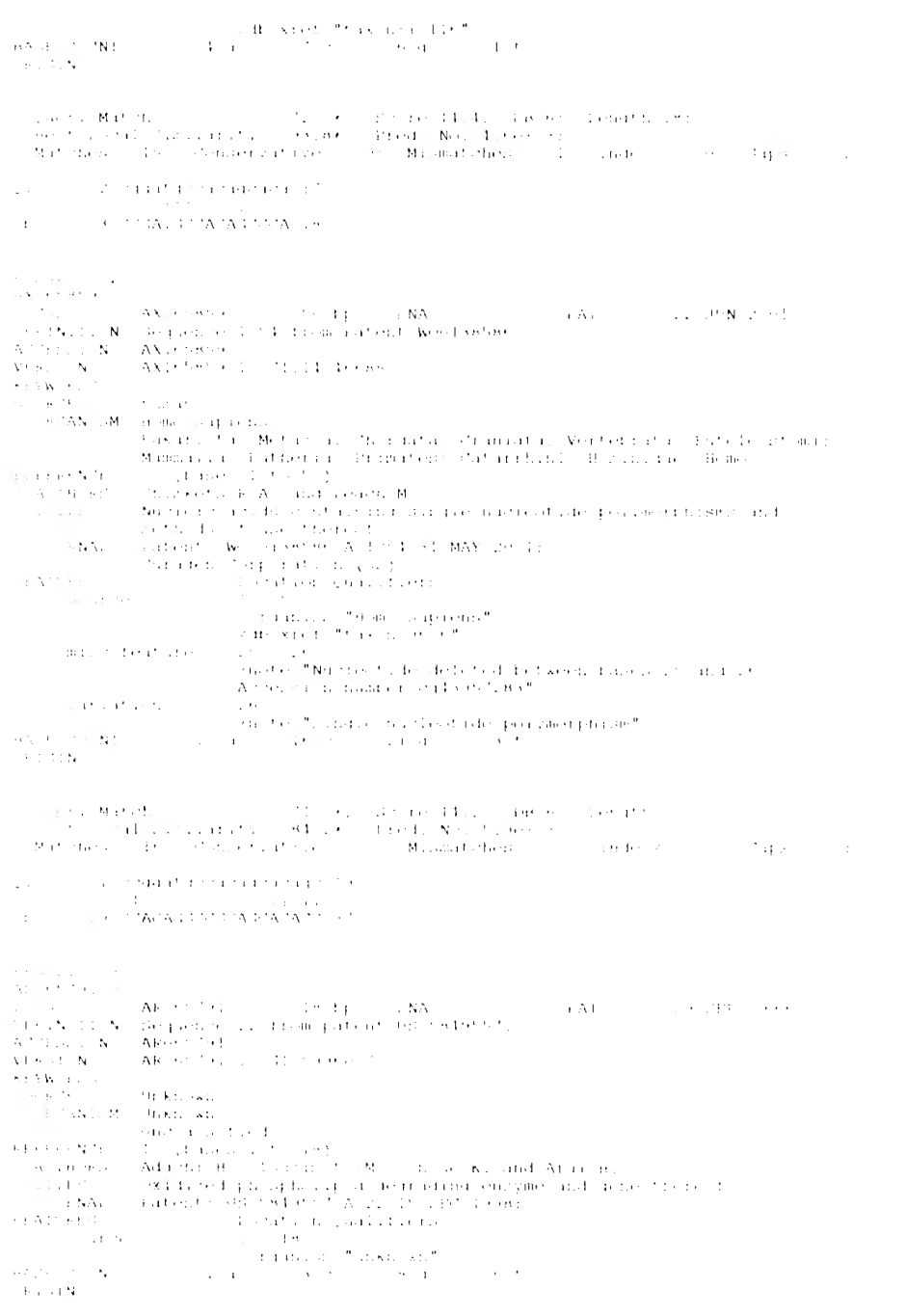
Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	14.4	72.0	18	6	AX174772	AX174772 Sequence
2	14.2	71.0	50	6	AX165859	AX165859 Sequence
c 3	13.8	69.0	18	6	AR065791	AR065791 Sequence
c 4	13.4	67.0	20	12	AR066263	AR066263 Synthetic
c 5	13.4	67.0	20	12	AR069272	AR069272 Synthetic
c 6	13.2	66.0	21	6	AX020518	AX020518 Sequence
c 7	13.2	66.0	24	6	AR107354	AR107354 Sequence
c 8	13.2	66.0	53	9	BDMFCRA13	L39016 Rat sulfot
c 9	13	65.0	50	6	AX076918	AX076918 Sequence
c 10	13	65.0	50	6	AX098266	AX098266 Sequence
11	12.8	64.0	34	6	AR045238	AR045238 Sequence
12	12.8	64.0	34	6	152290	152290 Sequence 31
c 13	12.6	63.0	21	6	A47967	A47967 Sequence 34
c 14	12.6	63.0	21	6	AX190975	AX190975 Sequence
c 15	12.6	63.0	23	6	AR041730	AR041730 Sequence
c 16	12.6	63.0	23	6	AR147381	AR147381 Sequence
17	12.6	63.0	30	6	A14390	A14390 1320 bp cDN
18	12.6	63.0	30	6	A48244	A48244 Sequence 25
19	12.6	63.0	30	6	147748	147748 Sequence 2
20	12.6	63.0	39	6	AR012246	AR012246 Sequence
21	12.6	63.0	39	6	114944	114944 Sequence 30
22	12.6	63.0	39	6	173664	173664 Sequence 30
23	12.6	63.0	40	6	A25552	A25552 174 bp cDN
24	12.6	63.0	40	6	147742	147742 Sequence 34
25	12.4	62.0	17	6	AX047544	AX047544 Sequence
c 26	12.4	62.0	20	6	AX191427	AX191427 Sequence
27	12.4	62.0	21	6	AR088918	AR088918 Sequence
c 28	12.4	62.0	31	6	188607	188607 Sequence 3
29	12.4	62.0	50	6	188606	188606 Sequence 2
30	12.4	62.0	54	6	A07526	A07526 Nucleotide
31	12.4	62.0	54	6	A12950	A12950 Nucleotide
32	12.2	61.0	18	6	E32484	E32484 Primer for
c 33	12.2	61.0	20	6	AX092801	AX092801 Sequence
c 34	12.2	61.0	21	6	AX164159	AX164159 Sequence
35	12.2	61.0	28	6	AX036321	AX036321 Sequence
36	12.2	61.0	38	6	AR024260	AR024260 Sequence
37	12.2	61.0	38	6	AR026928	AR026928 Sequence
38	12.2	61.0	38	6	AR045113	AR045113 Sequence
39	12.2	61.0	38	6	AR091814	AR091814 Sequence
40	12.2	61.0	38	6	E05404	E05404 DNA Sequenc
41	12.2	61.0	42	6	AR096930	AR096930 Sequence
42	12.2	61.0	47	6	A38847	A38847 Sequence 1
43	12.2	61.0	47	6	A38869	A38869 Sequence 1
44	12.2	61.0	47	6	A74057	A74057 Sequence 1
45	12.2	61.0	47	6	A74149	A74149 Sequence 1

ALIGNMENTS

```
RESULT 1
AX174772
LOCUS       AX174772               18 bp    DNA                 PAT       08-JAN-2001
DEFINITION  Sequence 5 from Patent WO0143759.
ACCESSION   AX174772
VERSION     AX174772.1  GI:14598292
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciuromorpha; Muroidea; Muridae;
            Rattus.
REFERENCE   1  (bases 1 to 18)
  AUTHORS   Campbell,C.A., Hadingham,S.J., Harrison,D.C. and Bertram,C.L.
  TITLE     New use
  JOURNAL   Patent: WO 0143759 A 5 21-JUN-2001;
            SMITHKLINE BEECHAM PLC (GB)
FEATURES             Location/Qualifiers
     source          1..18
                     /organism="Rattus norvegicus"
```



```
                and Suzuki Y.
  TITLE     A bac based sts content map spanning a 45 mb region of human
            chromosome 1p36 p36
  JOURNAL   Genomics 74 (1), 55-76 (2001)
  MEDLINE   21269192
REFERENCE   2  (bases 1 to 20)
  AUTHORS   Horii,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (04 AUG 2001) Akira Horii, Tohoku University School of
            Medicine, Molecular Pathology; 2-1 Seiryomachi, Aoba-ku, Sendai,
            Miyagi 980 8575, Japan (E-mail:horii@mail.cc.tohoku.ac.jp,
            Tel:81 22 717 8042, Fax:81 22 717 8047)
FEATURES             Location/Qualifiers
     source          1..20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     misc_feature    1..20
                     /note="forward primer for human STS sts-B125H4-F at 1p36
                     sts-B125H4-F obtained from clones B125H4, B23K15, B23L13,
                     B43H4, B123N19, B43A4, B268F17, Human BAC library
                     RPCI-11"
BASE COUNT         3 a         4 c         7 g         6 t
ORIGIN


  Query Match             67.0%;  Score 13.4;  DB 12;  Length 20;
  Best Local Similarity   93.3%;  Pred. No. 1.6e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 atgcgccacgcagcaagcca 24
           ||| ||||||||||| ||||
Db      20 ATGACACACACACACGCT 4


RESULT   6
AX020518/c
LOCUS       AX020518        21 bp    DNA             PAT       07-SEP-2000
DEFINITION  Sequence 18 from Patent WO9934016.
ACCESSION   AX020518
VERSION     AX020518.1  GI:10044208
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Vider,B.Z.
  TITLE     A method for identifying and characterizing cells and tissues
  JOURNAL   Patent: WO 9934016-A 18 08-JUL-1999;
            GENENA LTD (-) V VIDER BEN ZION (IL)
FEATURES             Location/Qualifiers
     source          1..21
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT         4 a         6 c         6 g         5 t
ORIGIN


  Query Match             66.0%;  Score 13.2;  DB 6;  Length 21;
  Best Local Similarity   83.3%;  Pred. No. 2.1e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 atgcgccacgcagcaagcca 24
           ||||||||||| || ||||
Db      19 GGATGCCACACGACTCCA 2


RESULT   7
AR107454/c
LOCUS       AR107454        24 bp    DNA             PAT       14-FEB-2001
DEFINITION  Sequence 7 from patent US 6110457.
ACCESSION   AR107454
VERSION     AR107454.1  GI:12822841
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Belshe,P.B. and Ray,R.
  TITLE     Live attenuated vaccines based on cp45 HPIV 4 strain and 2'-C to
            ensure attenuation in such vaccines
  JOURNAL   Patent: US 6110457 A 7 29-AUG-2000;
FEATURES             Location/Qualifiers
     source          1..24
                     /organism="unknown"
BASE COUNT         4 a         7 c         8 g         5 t
ORIGIN


  Query Match             66.0%;  Score 13.2;  DB 6;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 2.1e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 atgcgccacgcagcaagcca 24
           ||||||||||| || ||||
Db      21 GGATGCCACACGACTCCA 4


RESULT   8
HUMTCRAD3/c
LOCUS       HUMTCRAD3       53 bp    mRNA            PRI       21-AUG-1995
DEFINITION  Homo sapiens (C.A.LR03) rearranged T cell receptor delta chain
            (TCRAV19-TCRDJ1) mRNA, partial V region.
ACCESSION   L39616
VERSION     L39616.1  GI:945299
KEYWORDS    CDR3 region; T cell receptor alpha chain; T cell receptor beta,
            antigen recognition site; junctional region; rearranged variable
            region.
SOURCE      Homo sapiens (clone: C.A.LR03) colon cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 53)
  AUTHORS   Hoffmeier,W., Chowers,Y., Lamendo,A., Morzycka-Wroblewska,E. and
            Kaqnoff,M.F.
  TITLE     The delta T cell receptor repertoire in human colon and peripheral
            blood is oligoclonal irrespective of V region usage
  JOURNAL   J. Clin. Invest. 96 (2), 1108-1117 (1995)
  MEDLINE   95362812
COMMENT     Citation paper.
FEATURES             Location/Qualifiers
     source          1..53
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="C.A.LR03"
                     /tissue_type="colon"
     V_region        <1..>53
                     /gene="TCRAV19-TCRDJ1"
                     /standard_name="(V alpha 19/J delta 1)"
                     /note="putative"
     gene            1..53
                     /gene="TCRAV19-TCRDJ1"
     CDS             <1..>53
                     /gene="TCRAV19-TCRDJ1"
                     /note="This CDS feature is included to show the
                     translation of the corresponding V region. Presently
                     translation qualifiers on V region features are literal."
                     /codon_start=1
                     /protein_id="AAC41818.1"
                     /db_xref="GI:950497"
                     /translation="YFCAVRSGFGTGLTDKL"
BASE COUNT        12 a        14 c        12 g        15 t
ORIGIN


  Query Match             66.0%;  Score 13.2;  DB 9;  Length 53;
```

```
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Jannes,G., Rossau,R. and Van,H.H.
  TITLE     SIMULTANEOUS DETECTION, IDENTIFICATION AND DIFFERENTIATION OF
            SUBBACTERIAL TAXA USING A HYBRIDIZATION ASSAY
  JOURNAL   Patent: WO 9600298 A 44 04 JAN 1996;
            INNOGENETICS NV (BE)
COMMENT     other publication AU 29246/95 960119.
FEATURES             Location/Qualifiers
     source          1..21
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT         2 a        4 c        8 g        7 t
ORIGIN


  Query Match             63.0%;  Score 12.6;  DB 6;  Length 21;
  Best Local Similarity   78.9%;  Pred. No. 4.3e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 cgaatgccacagcacagcca 20
           ||||||  || ||| ||||
Db      18 CGCATGAGACTGCAGAGCA 5


RESULT  14
AX190775/c
LOCUS       AX190775        21 bp    DNA             PAT       10 AUG 2001
DEFINITION  Sequence 31 from Patent EP 1091064.
ACCESSION   AX190775
VERSION     AX190775.1  GI:15449620
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Jannes,G., Rossau,R. and Van Heerswynd,H.
  TITLE     Simultaneous detection, identification and differentiation of
            subbacterial taxa using a hybridization assay
  JOURNAL   Patent: EP 1091064 A 34 11 APR 2001;
            INNOGENETICS N.V. (BE)
FEATURES             Location/Qualifiers
     source          1..21
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT         2 a        4 c        8 g        7 t
ORIGIN


  Query Match             63.0%;  Score 12.6;  DB 6;  Length 21;
  Best Local Similarity   78.9%;  Pred. No. 4.3e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 cgaatgccacagcacagcca 20
           ||||||  || ||| ||||
Db      18 CGCATGAGACTGCAGAGCA 5


RESULT  15
AR041730/c
LOCUS       AR041730        23 bp    NA              PAT       29 SEP 1999
DEFINITION  Sequence 10 from patent US 5866404.
ACCESSION   AR041730
VERSION     AR041730.1  GI:5946016
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   Bradshaw,M.Suzanne, Bellekens,I.A. and Boddie,F.H.
  TITLE     Yeast bacteria shuttle vector
  JOURNAL   Patent: US 5866404 A 10 02 FEB 1999;
FEATURES             Location/Qualifiers
     source          1..23
                     /organism="unknown"
BASE COUNT         3 a        8 c        5 g        7 t
ORIGIN


  Query Match             63.0%;  Score 12.6;  DB 6;  Length 23;
  Best Local Similarity   78.9%;  Pred. No. 4.3e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 cgaatgccacagcacagcca 20
           ||||||  || ||| ||||
Db      20 CGCATGAGACTGCAGAGCA 7


Search completed: February  9, 2002, 01:43:52
Job time: 21881 sec
```

GenCore version 4.5

Run on: February 9, 2002, 01:43:52 ; Search time 2519.39 Seconds
(without alignments)
140.962 Million cell updates/sec

Title: US-10-006-430-52
Perfect score: 20
Sequence: 1 ctttccccaatgccaaccaa 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 147240 seqs, 82485875 residues

Total number of hits satisfying chosen parameters: 586446

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: gb_ba:*
2: gb_htc:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_mus:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_mus:*
35: em_htg_rod:*
36: em_htg_other:*

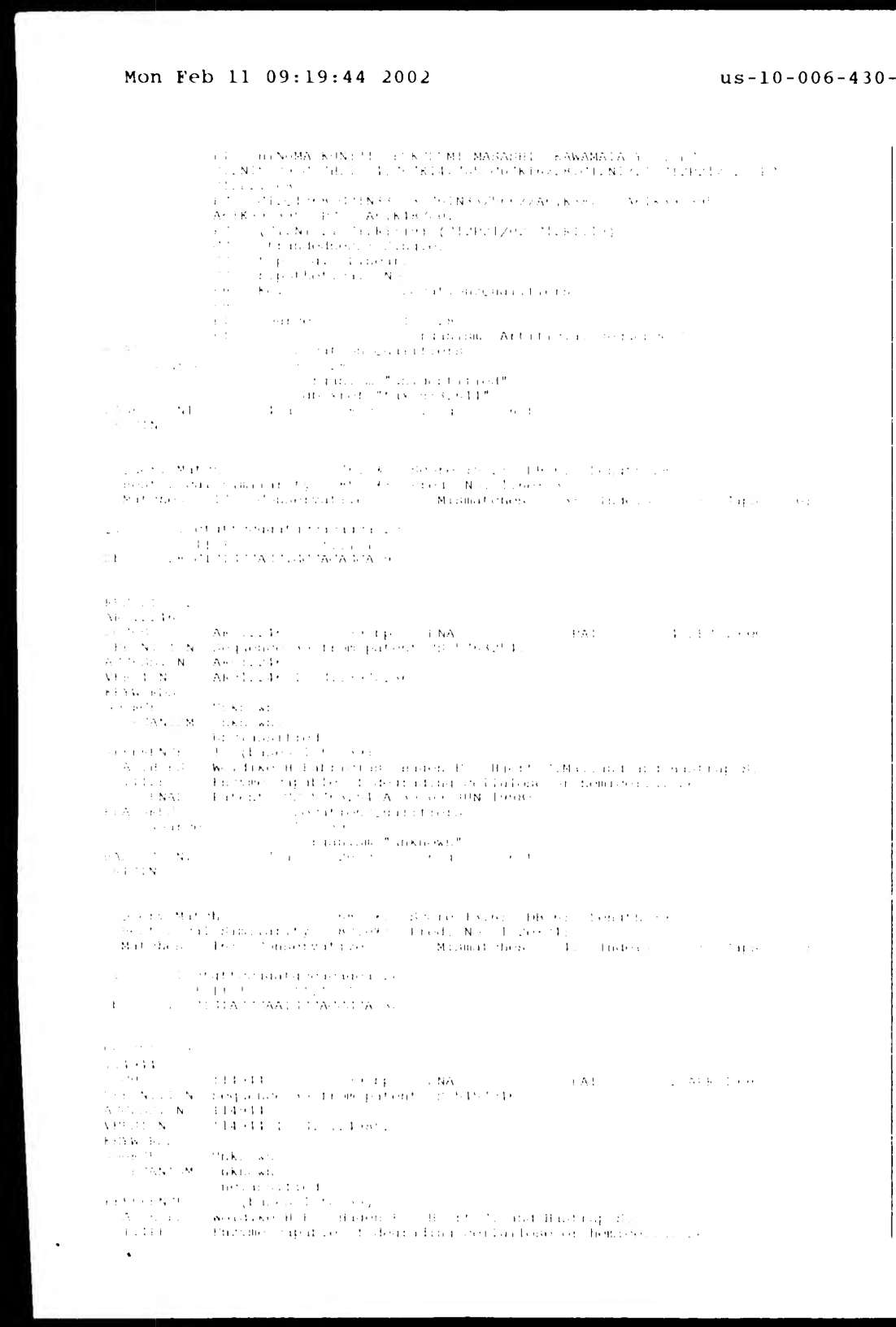
Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

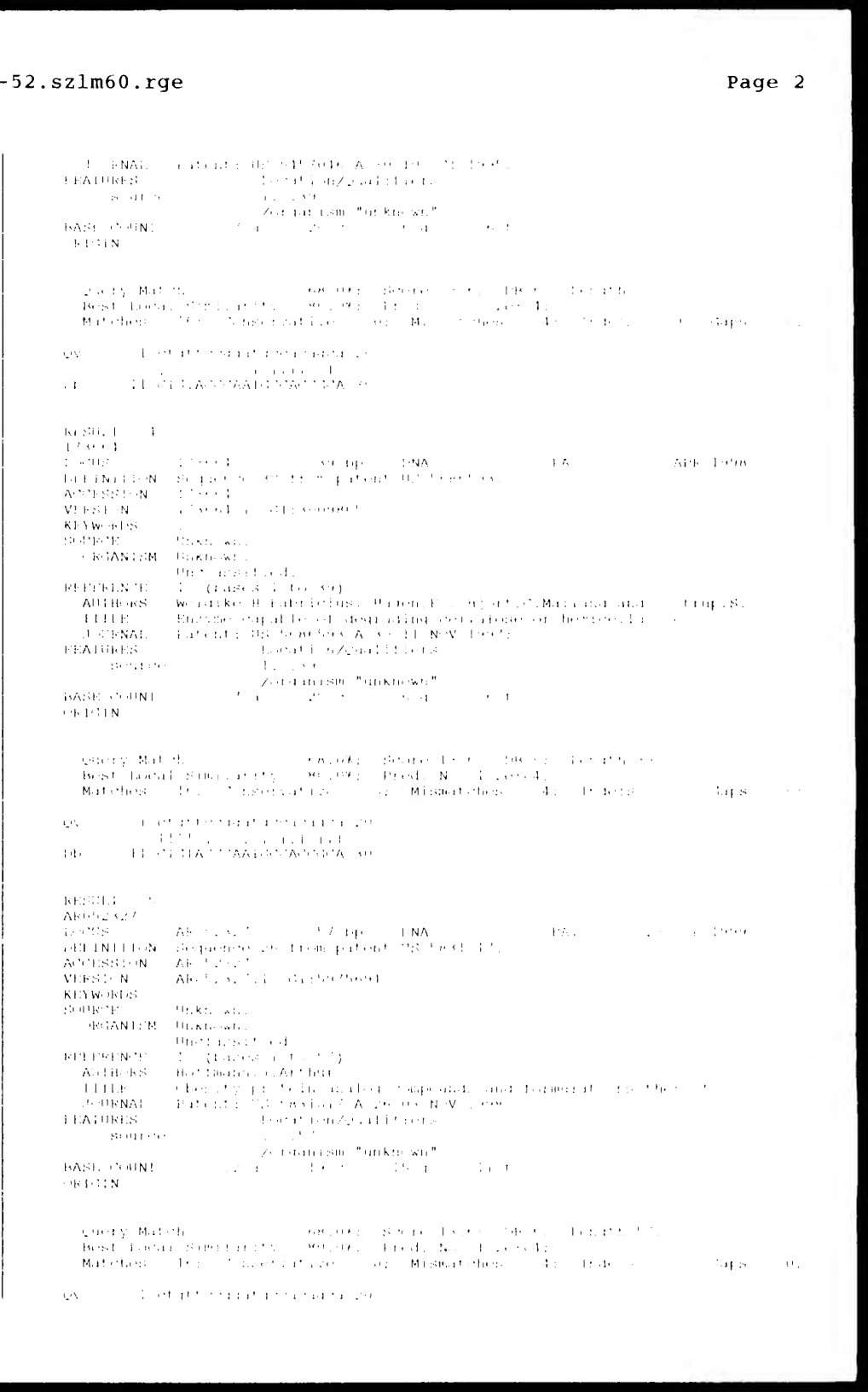
SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	15.2	76.0	28	6	E16002	E16002 PCR primer
2	13.6	68.0	39	6	AR012246	AR012246 Sequence
3	13.6	68.0	39	6	114944	114944 Sequence 39
4	13.6	68.0	39	6	173664	173664 Sequence 39
5	13.6	68.0	57	6	AR052427	AR052427 Sequence
6	13.4	67.0	20	12	AR068239	AR068239 Synthetic
7	13	65.0	18	6	AX174772	AX174772 Sequence
c 8	13	65.0	33	6	AR004366	AR004366 Sequence
c 9	13	65.0	33	6	AR097157	AR097157 Sequence
c 10	13	65.0	33	6	AR130655	AR130655 Sequence
c 11	13	65.0	33	6	182841	182841 Sequence 20
c 12	13	65.0	50	6	AX076918	AX076918 Sequence
c 13	13	65.0	50	6	AX098266	AX098266 Sequence
c 14	12.8	64.0	19	6	AR088415	AR088415 Sequence
15	12.8	64.0	37	6	A94879	A94879 Sequence 3
16	12.8	64.0	37	6	E36501	E36501 Varicella z
17	12.6	63.0	54	6	AX107494	AX107494 Sequence
18	12.6	63.0	54	6	AX107408	AX107408 Sequence
19	12.2	61.0	21	6	AR072123	AR072123 Sequence
c 20	12.2	61.0	21	6	AX096008	AX096008 Sequence
c 21	12.2	61.0	24	6	A10766	A10766 Escherichia
c 22	12.2	61.0	27	6	E04079	E04079 Primer for
23	12.2	61.0	29	6	AR053301	AR053301 Sequence
24	12.2	61.0	29	6	AR079507	AR079507 Sequence
25	12.2	61.0	29	6	AX164668	AX164668 Sequence
26	12.2	61.0	29	6	140920	140921 Sequence 35
27	12.2	61.0	45	6	AX107691	AX107692 Sequence
28	12.2	61.0	49	6	AX107695	AX107695 Sequence
29	12.2	61.0	51	6	AX165351	AX165351 Sequence
c 30	12.2	61.0	51	9	HSA224253	AL224253 Homo sapi
c 31	12.2	61.0	58	1	AF195104	AF195104 Silicibac
32	12	60.0	20	6	AX082432	AX082432 Sequence
33	12	60.0	20	6	AX084783	AX084783 Sequence
34	12	60.0	23	6	124871	124871 Sequence 3
35	12	60.0	30	6	AR109718	AR109718 Sequence
c 36	12	60.0	30	10	MMBNP114A	X95954 M.musculus
37	12	60.0	31	6	AX055854	AX055854 Sequence
c 38	12	60.0	32	6	AX003253	AX003253 Sequence
c 39	12	60.0	32	6	AX004567	AX004567 Sequence
c 40	12	60.0	32	6	AX004587	AX004587 Sequence
c 41	12	60.0	32	6	AX018569	AX018569 Sequence
c 42	12	60.0	32	6	AX018645	AX018645 Sequence
c 43	12	60.0	32	6	AX023698	AX023698 Sequence
c 44	12	60.0	32	6	AX046658	AX046658 Sequence
c 45	12	60.0	39	10	MM027225	L27225 Mus musculu

ALIGNMENTS

RESULT 1
E16002/c
LOCUS E16002 28 bp DNA PAT 28-JUL-1998
DEFINITION PCR primer for guiding human G protein coupled receptor beta
ACCESSION E16002
VERSION E16002.1 GI:5710685
KEYWORDS JP 19981127289 A/6.
SOURCE unidentified.
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 28)
AUTHORS Himmai,K., Tokuzumi,M. and Kawai,Y.
TITLE NEW G PROTEIN CONJUGATE TYPE RECEPTOR PROTEIN ASSAY
JOURNAL Patent: JP 19981127289 A 6 19-MAY-1998;
FAKEDA CHEM IND LTD
COMMENT OS None
OC Artificial sequences.
PN JP 19981127289 A/6
PD 19-MAY-1998
PF 29-OCT-1996 JP 19962868





```
          1 [ ]]] ] ]]]].-]]]]
Qy      21 'AGCTCCAGCTGCCACAGCA 42


RESULT   6
AR068239
LOCUS       AR068239          20 bp    DNA             SYN       08-AUG-2001
DEFINITION  Synthetic construct DNA, reverse primer for human STS
            sts:stGDB:198890 at 1p36.
ACCESSION   AR068239
VERSION     AR068239.1  GI:15122044
KEYWORDS    .
SOURCE      synthetic construct DNA
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Chen,Y.Z., Bayashi,Y., Wu,I.C., Takaoka,E., Maekawa,K.,
            Watanabe,N., Tsurowa,L., Rosoki,E., Arai,Y., Mizushima,H.,
            Morohashi,A., Ohita,M., Nakanawara,A., Lin,S., Hoshi,M., Horii,A.
            and Sendi,E.
  TITLE     A bac based sts content map spanning a 45 mb region of human
            chromosome 1p36 p36
  JOURNAL   Genomics. 74 (1), 55 Jn (2001)
  MEDLINE   21265702
REFERENCE   2  (bases 1 to 20)
  AUTHORS   Horii,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (04 Au   -01) Akira Horii, Tohoku University School of
            Medicine, Molecular Pathology; 2-1 Seiryomachi, Aoba-ku, Sendai,
            Miyagi 980-8575, Japan (E-mail:horii@mail.cc.tohoku.ac.jp,
            Tel:81 22 717 8042, Fax:81 22 717 8047)
FEATURES             Location/Qualifiers
     source          1..20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     misc_feature    1..20
                     /note="reverse primer for human STS sts:stGDB:198890 at
                     1p36
                     sts stGDB:198890 obtained from clones B142I12, B061M21,
                     B146J5, B299A22, B243I3, B471L1, B151M24, Human BAC
                     library RPCI-11"
BASE COUNT         3 a        4 c        7 g        6 t
ORIGIN


  Query Match           67.0%;  Score 13.4;  DB 12;  Length 20;
  Best Local Similarity   93.8%;  Prod. No. 1.6e-04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 tgttccggatgccaac 16
           ||||| ||||||||
Db       1 TGTTCTGGATGCCAC 15


RESULT   7
AX174772
LOCUS       AX174772          18 bp    DNA             PAT       03-JUL-2001
DEFINITION  Sequence 5 from Patent WO0144759.
ACCESSION   AX174772
VERSION     AX174772.1  GI:14508202
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 18)
  AUTHORS   Campbell,C.A., Burningham,S.J., Harrison,D.C. and Berzieu,G.J.
  TITLE     New use
  JOURNAL   Patent: WO 0144759-A 5 21-JUN 2001;
            SMITHKLINE BEECHAM PLC (GB)
FEATURES             Location/Qualifiers
     source          1..18
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
BASE COUNT         4 a        7 c        6 g        1 t
ORIGIN


  Query Match           65.0%;  Score 13;  DB 6;  Length 18;
  Best Local Similarity   100.0%;  Prod. No. 2.6e-04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 cgaatgccacagc 19
           |||||||||||||
Db       4 CGGATGCCACAGC 16


RESULT   8
AR004466/c
LOCUS       AR004466          33 bp    DNA             PAT       04-DEC-1998
DEFINITION  Sequence 20 from patent US 5747244
ACCESSION   AR004466
VERSION     AR004466.1  GI:956245
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Sheridan,P., Chang,C., Ronning,J. and Gilbert,M.S.
  TITLE     Nucleic acid probes immobilized on polystyrene surfaces
  JOURNAL   Patent: US 5747244 A 20-05-MAY-1998;
FEATURES             Location/Qualifiers
     source          1..33
                     /organism="unknown"
BASE COUNT         6 a        5 c       12 g        6 t        4 others
ORIGIN


  Query Match           65.0%;  Score 13;  DB 6;  Length 33;
  Best Local Similarity   76.5%;  Prod. No. 2.6e-04;
  Matches   13;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ctgttccggatgccaca 17
           :|: ||||||| |||||
Db      24 STRCTGGGATGCCACA 8


RESULT   9
AR097157/c
LOCUS       AR097157          33 bp    DNA             PAT       14-FEB-2001
DEFINITION  Sequence 95 from patent US 6071693.
ACCESSION   AR097157
VERSION     AR097157.1  GI:12805887
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Cha,F., Beall,E., Irvine,B., Kolberg,J. and Urdea,M.S.
  TITLE     HCV genomic sequences for diagnostics and therapeutics
  JOURNAL   Patent: US 6071693-A 95 06-JUN 2000;
FEATURES             Location/Qualifiers
     source          1..33
                     /organism="unknown"
BASE COUNT         6 a        5 c       12 g        6 t        4 others
ORIGIN


  Query Match           65.0%;  Score 13;  DB 6;  Length 33;
  Best Local Similarity   76.5%;  Prod. No. 2.6e-04;
  Matches   13;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ctgttccggatgccaca 17
           :|: ||||||| |||||
Db      24 STRCTGGGATGCCACA 8
```

```
RESULT  4
AX000841/c
LOCUS       AX000841               97 bp    DNA             PAT       07-APR-2000
DEFINITION  Sequence 4 from patent US 6000085.
ACCESSION   AX000841
VERSION     AX000841.1  GI:3455678
KEYWORDS    .
SOURCE      Mouse sp.
  ORGANISM  Mouse sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciuridae; Sciurus.
REFERENCE   1  (bases 1 to 97)
  AUTHORS   Lee,S.H., Lefkowitz,S.V., Petty,W.I., III and Skeeter, T.M.H.
  TITLE     Specific mutations of map kinase V (MKK4) in human lung cell lines
            identify it as a tumor suppressor in various types of cancer
  JOURNAL   Patent: US 6000085-A 1 24-NOV-1999

RESULT  3
AX000840/c
LOCUS       AX000840               97 bp    DNA             PAT       07-APR-2000
DEFINITION  Sequence 3 from patent US 6000085.
ACCESSION   AX000840
VERSION     AX000840.1  GI:3455677
KEYWORDS    .
SOURCE      Synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 97)
  AUTHORS   Lee,S.H., Lefkowitz,S.V., Petty,W.I., III and Skeeter, T.M.H.
  TITLE     Specific mutations of map kinase V (MKK4) in human lung cell lines
            identify it as a tumor suppressor in various types of cancer
  JOURNAL   Patent: US 6000085-A 1 24-NOV-1999

RESULT  2
AX000839/c
LOCUS       AX000839               97 bp    DNA             PAT       07-APR-2000
DEFINITION  Sequence 2 from patent US 6000085.
ACCESSION   AX000839
VERSION     AX000839.1  GI:3455676
KEYWORDS    .
SOURCE      Mouse sp.
  ORGANISM  Mouse sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciuridae; Sciurus.
REFERENCE   1  (bases 1 to 97)
  AUTHORS   Lee,S.H., Lefkowitz,S.V., Petty,W.I., III and Skeeter, T.M.H.
  TITLE     Specific mutations of map kinase V (MKK4) in human lung cell lines
            identify it as a tumor suppressor in various types of cancer
  JOURNAL   Patent: US 6000085-A 1 24-NOV-1999

RESULT  1
AX000838/c
LOCUS       AX000838               97 bp    DNA             PAT       07-APR-2000
DEFINITION  Sequence 1 from patent US 6000085.
ACCESSION   AX000838
VERSION     AX000838.1  GI:3455675
KEYWORDS    .
SOURCE      Mouse sp.
  ORGANISM  Mouse sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciuridae; Sciurus.
REFERENCE   1  (bases 1 to 97)
  AUTHORS   Lee,S.H., Lefkowitz,S.V., Petty,W.I., III and Skeeter, T.M.H.
  TITLE     Specific mutations of map kinase V (MKK4) in human lung cell lines
            identify it as a tumor suppressor in various types of cancer
  JOURNAL   Patent: US 6000085-A 1 24-NOV-1999
```

```
FEATURE             Location/Qualifiers
     source         1..39
                    /organism="unknown"
BASE COUNT        7 a      5 c      4 g      8 t
ORIGIN


  Query Match             64.0%;  Score 12.8;  DB 6;  Length 39;
  Best Local Similarity   87.5%;  Pred. No. 3.4e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 ttcccgatgcccaaca 20
           || ||||||||||| |
Db      16 tcacgatgccacacaa 1


RESULT  15
A04879
LOCUS       A04879             47 bp    DNA             PAT       26-JAN-2000
DEFINITION  Sequence 4 from Patent EP0627470.
ACCESSION   A04879
VERSION     A04879.1  GI:6779076
KEYWORDS
SOURCE      Human herpesvirus 4.
  ORGANISM  Human herpesvirus 4
            Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
            Alphaherpesvirinae; Varicellovirus.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Giesendorf,B.D. and Liebkamm,M.D.
  TITLE     Immunoreactive protein VP26 of Varicella Zoster Virus (VZV) and its
            diagnostic use
  JOURNAL   Patent: EP 0627470 A 4 07-JUL-1999;
            DADE BEHRING MARBURG GMBH (DE)
FEATURES             Location/Qualifiers
     source          1..47
                     /organism="Human herpesvirus 4"
                     /strain="ELLEN"
                     /db_xref="taxon:10376"
                     /map="ORF26"
BASE COUNT        9 a     11 c      7 g      8 t
ORIGIN


  Query Match             64.0%;  Score 12.8;  DB 6;  Length 47;
  Best Local Similarity   87.5%;  Pred. No. 3.4e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 tttcccgatgcccaac 19
           |||||||| |||| |
Db       5 tttcccgatgacacaa 20


Search completed: February  9, 2002, 01:43:54
Job time: 21883 sec
```

GenCore version 4.5

Run on: February 9, 2002, 01:44:55 ; Search time 2519.49 Seconds
(without alignments)
130.962 Million cell updates/sec

Title: US-10-006-430-54
Perfect score: 20
Sequence: 1 qtcaaaqqapqetqttcaqqqa 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 823859755 residues

Total number of hits satisfying chosen parameters: 986436

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_phg:*
8: gb_ph:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_mus:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	13.8	69.0	30	6	AR026490	AR026490 Sequence 15
c 2	13.2	66.0	41	6	AR032338	AR032338 Sequence 1
3	13	65.0	18	6	AR026492	AR026492 Sequence 17
c 4	12.8	64.0	30	6	AX027301	AX027301 Sequence 1
c 5	12.8	64.0	40	6	AR079802	AR079802 Sequence 1
c 6	12.8	64.0	40	6	AR081332	AR081332 Sequence 1
7	12.8	64.0	51	6	AX161845	AX161845 Sequence 1
8	12.8	64.0	51	9	BSZ95085	Z95085 H.sapiens m
9	12.6	63.0	45	8	S65078	S65078 VL7/HBG2 al
10	12.6	63.0	48	8	S65072	S65072 VL7/KL2 al
11	12.6	63.0	51	6	AX161851	AX161851 Sequence 1
12	12.6	63.0	51	6	AX161852	AX161852 Sequence 1
13	12.6	63.0	57	8	S65071	S65071 VL7/HBG2 al
c 14	12.6	63.0	60	6	E43730	E43730 Mutated lox
15	12.6	63.0	60	6	E43731	E43731 Mutated lox
c 16	12.6	63.0	60	6	E43732	E43732 Mutated lox
17	12.6	63.0	60	6	E43733	E43733 Mutated lox
18	12.4	62.0	29	6	AR07761	AR07761 Sequence 16
19	12.4	62.0	29	6	AX022725	AX022725 Sequence 1
c 20	12.2	61.0	20	6	A47383	A47383 Sequence 11
c 21	12.2	61.0	21	6	A47382	A47382 Sequence 10
c 22	12.2	61.0	22	6	A47381	A47381 Sequence 9
c 23	12.2	61.0	23	6	A47380	A47380 Sequence 8
c 24	12.2	61.0	24	6	A47379	A47379 Sequence 7
c 25	12.2	61.0	25	6	A47378	A47378 Sequence 6
c 26	12.2	61.0	25	6	A47397	A47397 Sequence 25
c 27	12.2	61.0	25	6	A47398	A47398 Sequence 26
c 28	12.2	61.0	25	6	A47399	A47399 Sequence 27
c 29	12.2	61.0	25	6	A47400	A47400 Sequence 28
c 30	12.2	61.0	25	6	A47401	A47401 Sequence 29
c 31	12.2	61.0	25	6	A47402	A47402 Sequence 30
c 32	12.2	61.0	25	6	A47403	A47403 Sequence 31
c 33	12.2	61.0	25	6	A47404	A47404 Sequence 32
c 34	12.2	61.0	25	6	A47405	A47405 Sequence 33
c 35	12.2	61.0	25	6	A47406	A47406 Sequence 34
c 36	12.2	61.0	25	6	A47407	A47407 Sequence 35
c 37	12.2	61.0	25	6	A47408	A47408 Sequence 36
c 38	12.2	61.0	25	6	A47409	A47409 Sequence 37
c 39	12.2	61.0	25	6	A47410	A47410 Sequence 38
c 40	12.2	61.0	25	6	A47411	A47411 Sequence 39
c 41	12.2	61.0	25	6	A47412	A47412 Sequence 40
c 42	12.2	61.0	26	6	A47377	A47377 Sequence 5
c 43	12.2	61.0	27	6	A47376	A47376 Sequence 4
c 44	12.2	61.0	28	6	A47375	A47375 Sequence 3
c 45	12.2	61.0	29	6	A47374	A47374 Sequence 2

ALIGNMENTS

RESULT 1
AR026490
LOCUS AR026490 30 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 15 from patent US 5856096.
ACCESSION AR026490
VERSION AR026490.1 GI:5937430
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Windle,B.E., Qin,M., Chen,S., Fletcher,J.M. and Maine,I.
TITLE Rapid and sensitive assays for detecting and distinguishing between processive and non processive telomerase activities
JOURNAL Patent: US 5856096 A 15 05-JAN-1999;
FEATURES Location/Qualifiers
source 1..30
/organism="unknown"
BASE COUNT 8 a 11 c 7 g 4 t
ORIGIN

```
  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy [illegible]
Db [illegible] ACATTCATTTCA[illegible]

RESULT [illegible]
[illegible]
LOCUS [illegible] DNA PAT [illegible]
DEFINITION [illegible]
ACCESSION [illegible]
VERSION [illegible]
KEYWORDS [illegible]
SOURCE Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE 1 (bases 1 to [illegible])
  AUTHORS [illegible]
  TITLE [illegible]
  JOURNAL Patent: [illegible]
FEATURES Location/Qualifiers
     source 1..[illegible]
            /organism="unknown"
BASE COUNT [illegible]
ORIGIN

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy [illegible]
Db [illegible] ACATTCATTTCA[illegible]

RESULT [illegible]
[illegible]
LOCUS AR[illegible] [illegible] bp DNA PAT [illegible] SEP [illegible]
DEFINITION Sequence [illegible] from patent US [illegible]
ACCESSION AR[illegible]
VERSION AR[illegible]
KEYWORDS [illegible]
SOURCE Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE 1 (bases 1 to [illegible])
  AUTHORS [illegible]
  TITLE [illegible]
  JOURNAL Patent: [illegible] JAN [illegible]
FEATURES Location/Qualifiers
     source 1..[illegible]
            /organism="unknown"
BASE COUNT [illegible]
ORIGIN

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy [illegible]
Db [illegible] ACATTCATT[illegible]

RESULT [illegible]

AX[illegible]
LOCUS AX[illegible] [illegible] bp DNA PAT [illegible] SEP [illegible]
DEFINITION Sequence 1 from Patent FR[illegible]
ACCESSION AX[illegible]
VERSION AX[illegible] GI:[illegible]
KEYWORDS .
SOURCE synthetic construct.
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE 1 (bases 1 to [illegible])
  AUTHORS Beaudet,A., Vaudry,H., Tonkes,J., Chartrel,N. and [illegible]
  JOURNAL Patent: FR [illegible] A 1 [illegible]
            INST NAT SANTE RECH MED (FR)
FEATURES Location/Qualifiers
     source 1..[illegible]
            /organism="synthetic construct"
            /db_xref="taxon:[illegible]"
            /note="amorce [illegible]"
BASE COUNT [illegible]
ORIGIN

  Query Match [illegible] Score [illegible] DB [illegible] Length [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy 4 [illegible]
Db [illegible] CACTCATT[illegible]

RESULT 5
AR[illegible]
LOCUS AR[illegible] [illegible] bp DNA PAT [illegible] 2000
DEFINITION Sequence [illegible] from patent US [illegible]
ACCESSION AR[illegible]
VERSION AR[illegible] GI:[illegible]
KEYWORDS .
SOURCE Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE 1 (bases 1 to [illegible])
  AUTHORS Ballakis,C.N. and Felber,B.K.
  TITLE Method of eliminating inhibitory/instability regions [illegible] mRNA
  JOURNAL Patent: US [illegible] A [illegible]
FEATURES Location/Qualifiers
     source 1..[illegible]
            /organism="unknown"
BASE COUNT [illegible]
ORIGIN

  Query Match [illegible] Score [illegible] DB [illegible] Length [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches 14; Conservative 0; Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy 1 [illegible]
Db [illegible] GTTCAT[illegible]

RESULT 6
AR[illegible]
LOCUS AR[illegible] [illegible] bp DNA PAT [illegible] 2000
DEFINITION Sequence [illegible] from patent US [illegible]
ACCESSION AR[illegible]
VERSION AR[illegible] GI:[illegible]
KEYWORDS .
SOURCE Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE 1 (bases 1 to [illegible])
  AUTHORS Pavlakis,G.N. and Felber,B.K.
```

```
  Title      Nucleic acid constructs containing HIV genes with mutated
             inhibitory/instability re  ns and methods of using same
  Journal    Patent: US 5,612,500 A 18-Mar-1997
FEATURES             Location/Qualifiers
     source          1..40
                     /organism "unknown"
BASE COUNT        12 a      10 c      12 g       6 t
ORIGIN


  Query Match           64.0%;  Score 12.8;  DB 6;  Length 40;
  Best Local Similarity   87.5%;  Pred. No. 7.7e-04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 atgcaagcagctgctc 16
           |||||||||||||||
Db      11 ATGCAGGCAGCTGCTC 26

RESULT   7
AX616045
LOCUS       AX616045          51 bp    DNA             PAT       22-JUN-2001
DEFINITION  Sequence 5164 from Patent WO0140521.
ACCESSION   AX616045
VERSION     AX616045.1  GI:14544116
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 5164 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    25
                     /note="1 of 2 allelic variants (5164 is other entry)
                     Accession number gi|4390467|"
BASE COUNT        13 a      12 c       8 g      18 t
ORIGIN


  Query Match           64.0%;  Score 12.8;  DB 6;  Length 51;
  Best Local Similarity   87.5%;  Pred. No. 7.4e-04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 tgcaagcagctgctc 17
           |||||| ||||| |||
Db      10 TGCAGGAAGCTGCTC 24

RESULT   8
BE920685
LOCUS       BE920685          51 bp    mRNA            PRI       25-NOV-1998
DEFINITION  H.sapiens mRNA for T cell receptor beta chain V-(D) junctional CDR3
            region (TCRBV2 BJ2S2).
ACCESSION   Z95085
VERSION     Z95085.1  GI:928440
KEYWORDS    joining region; junctional region; T cell receptor; T cell receptor
            beta chain; variable region.
SOURCE      human
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Bachelez,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-1997) Bachelez H., Hopital Saint Louis, 1, avenue
            Claude Vellefaux, 75475 Paris cedex 10, FRANCE
REFERENCE   2  (bases 1 to 51)
  AUTHORS   Bachelez,H., Hadida,F., Parizot,C., Flageul,B., Perina,M.,
            Imbertret,L., Debree,P. and Autran,B.
  TITLE     Oligoclonal expansion of HIV-specific T cells in skin and
            in the skin of HIV-1 infected patients with cutaneous
            pseudolymphoma
  JOURNAL   J. Clin. Invest. 101 (11), 2500-2516 (1998)
  MEDLINE   98282286
COMMENT     On Nov 25, 1998 this sequence version replaced gi:1172...
FEATURES             Location/Qualifiers
     source          1..51
                     /organism "Homo sapiens"
                     /db_xref "taxon:9606"
                     /rearranged
     V_region        1..45
                     /note "TCRBV2 BJ2S2 transcribed region"
                     /product "T-cell receptor beta chain, variable region"
BASE COUNT        7 a      12 c      21 g      11 t
ORIGIN


  Query Match           64.0%;  Score 12.8;  DB 6;  Length 51;
  Best Local Similarity   87.5%;  Pred. No. 7.4e-04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 atgcaagcagctgctc 16
           |||| ||||||||||
Db      40 GTACAGGCAGCTGCTC 45

RESULT   9
S65078
LOCUS       S65078            45 bp    mRNA            INV       24-NOV-1993
DEFINITION  Vdl/RPD2 alpha peptide (alternatively spliced) [rat, AMPC, lymnaea
            stagnalis snails, central nervous system, mRNA Partial, 45 nt].
ACCESSION   S65078
VERSION     S65078.1  GI:410546
KEYWORDS    .
SOURCE      great pond snail central nervous system.
  ORGANISM  Lymnaea stagnalis
            Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata;
            Basommatophora; Lymnaeidae; Lymnaea.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Bouwer,J., Van Kesteren,R.E., Van Heerikhuizen,H., Geraerts,W.P.,
            Veenstra,J., Smit,A.B. and Jonker,A.
  TITLE     Alternative splicing generates diversity of Vdl/kd; alpha peptid
            in the central nervous system of Lymnaea stagnalis
  JOURNAL   Cell. Mol. Neurobiol. 13 (2), 123-136 (1993)
  MEDLINE   93351194
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 136946] from the original journal article.
            This sequence comes from Fig. 5.
FEATURES             Location/Qualifiers
     source          1..45
                     /organism "Lymnaea stagnalis"
                     /db_xref "taxon:6523"
     gene            1..45
                     /partial
                     /gene "Vdl/RPD2 alpha peptide"
     CDS             1..45
                     /partial
                     /gene "Vdl/RPD2 alpha peptide"
                     /note "this sequence comes from Fig. 5"
                     /codon_start=1
                     /product "Vdl/RPD2 alpha peptide"
                     /protein_id "AAB27755.1"
                     /db_xref "GI:410547"
                     /translation "DMRCQAHERNCPGEN"
BASE COUNT        12 a      11 c       9 g      13 t
ORIGIN
```

```
REFERENCE   1  (bases 1 to 57)
  AUTHORS   Boogerd,J., Van Kesteren,P.E., Van Heerikhuizen,H., Geraerts,W.P.,
            Veenstra,J., Smit,A.B. and Bosse,J.
  TITLE     Alternative splicing generates diversity of VDI/RPD2 alpha peptides
            in the central nervous system of Lymnaea stagnalis
  JOURNAL   Cell. Mol. Neurobiol. 13 (2), 143-146 (1993)
  MEDLINE   93351194
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 146944] from the original journal article.
            This sequence comes from Fig. 5.
FEATURES             Location/Qualifiers
     source          1..57
                     /organism="Lymnaea stagnalis"
                     /db_xref="taxon:6523"
     gene            1..57
                     /partial
                     /gene="VDI/RPD2 alpha peptide"
     CDS             1..57
                     /partial
                     /gene="VDI/RPD2 alpha peptide"
                     /note="This sequence comes from Fig. 5"
                     /codon_start=1
                     /product="VDI/PPD2 alpha peptide"
                     /protein_id="AAB27753.1"
                     /db_xref="GI:410544"
                     /translation="MMCLASRCLRHPPENPCGEN"
BASE COUNT        14 a      13 c      15 g      15 t
ORIGIN


  Query Match             63.0%;  Score 12.6;  DB 6;  Length 57;
  Best Local Similarity   78.9%;  Pred. No. 9.2e-04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 tacacgaagctattccgga 20
           | ||| || | |||||||
Db      34 TATACGAACGTAATCCGGA 54


RESULT  14
E33740
LOCUS       E33740          60 bp    DNA              PAT       07-FEB-2001
DEFINITION  Mutated loxP sequence and application thereof.
ACCESSION   E33740
VERSION     E33740.1  GI:13018779
KEYWORDS    JP 11991196880-A/57.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Izumi,S.K.I.T.
  TITLE     Mutated loxP sequence and application thereof
  JOURNAL   Patent: JP 11991196880-A/57 27-JUL-1999;
            SUMITOMO PHARMACEUT CO LTD
COMMENT     OS   Unidentified
            PN   JP 11991196880-A/57
            PD   27-JUL-1999
            PF   28-SEP-1998 JP 19982273150
            PR
            PI   IZUMI SATIO;KEIJI TANAKA
            DI   Ç¼ÇÃ ÇÝÇÄÇÍ;ÇÃÇ¼ÇÃ ÇÜÇÂÇÈ
            DC   (Ç¼ÇNIÇ¾, Ç¾ÇiÇF Ç¾);Ç¼ÇÝÇÄÇÍ ÇÜÇNÇ 2Ç¼ÇNIÇ 1,61Ç¼,)
            CC
            FH   Key             Location/Qualifiers
            FT   source          1..60
            FT                   /organism="Unidentified".
FEATURES             Location/Qualifiers
     source          1..60
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        19 a      11 c      10 g      20 t
ORIGIN


  Query Match             63.0%;  Score 12.6;  DB 6;  Length 60;
  Best Local Similarity   78.9%;  Pred. No. 9.2e-04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 tacacgaagctattccgga 20
           | ||| || | |||||||
Db      41 TATACGAACGTAATCCGGA 59


Search completed: February  9, 2002, 01:19:55
Job time: 21884 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Nucleic - nucleic search, using sw model

Run on: February 9, 2002, 01:41:51 ; Search time 2519.39 Seconds
(without alignments)
130.962 Million cell updates/sec

Title: US-10-006-430-60
Perfect score: 20
Sequence: 1 aaggtaccgaaaattcaaga 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.

Searched: 1422140 seqs, 824858075 residues

Total number of hits satisfying chosen parameters: 586436

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	14.2	71.0	51	6	AX160629	AX160629 Sequence
2	14.2	71.0	51	6	AX160630	AX160630 Sequence
c 3	13.6	68.0	55	10	AF357367	AF357367 Mus musc
4	13.2	66.0	24	6	A49142	A49142 Sequence 4
5	13	65.0	18	12	AB069260	AB069260 Synthetic
6	12.8	64.0	19	6	A60406	A60406 Sequence 1
7	12.8	64.0	21	6	AX095468	AX095468 Sequence
8	12.8	64.0	35	6	AR061541	AR061541 Sequence
9	12.8	64.0	35	6	AR108440	AR108440 Sequence
10	12.8	64.0	35	6	116397	116397 Sequence 22
11	12.8	64.0	35	6	166883	166883 Sequence 22
12	12.8	64.0	35	6	184977	184977 Sequence 22
13	12.8	64.0	51	6	AX162429	AX162429 Sequence
14	12.8	64.0	51	6	AX162430	AX162430 Sequence
c 15	12.6	63.0	31	6	AR121249	AR121249 Sequence
c 16	12.4	62.0	30	6	A47694	A47694 Sequence 2
c 17	12.4	62.0	40	6	AX025232	AX025232 Sequence
c 18	12.2	61.0	19	6	AX141716	AX141716 Sequence
c 19	12.2	61.0	20	6	AR027966	AR027966 Sequence
20	12.2	61.0	20	6	AR027969	AR027969 Sequence
c 21	12.2	61.0	20	6	AR047886	AR047886 Sequence
22	12.2	61.0	20	6	AR047889	AR047889 Sequence
c 23	12.2	61.0	20	6	AR060953	AR060953 Sequence
24	12.2	61.0	20	6	AR060956	AR060956 Sequence
c 25	12.2	61.0	20	6	AR064357	AR064357 Sequence
26	12.2	61.0	20	6	AR064360	AR064360 Sequence
c 27	12.2	61.0	20	6	AR070632	AR070632 Sequence
28	12.2	61.0	20	6	AR070635	AR070635 Sequence
c 29	12.2	61.0	20	6	AR079953	AR079953 Sequence
30	12.2	61.0	20	6	AR079956	AR079956 Sequence
c 31	12.2	61.0	20	6	AR095059	AR095059 Sequence
c 32	12.2	61.0	45	6	AX127327	AX127327 Sequence
c 33	12.2	61.0	45	6	AX127328	AX127328 Sequence
c 34	12.2	61.0	51	6	AX165568	AX165568 Sequence
35	12.2	61.0	59	9	BSA241953	A1.1 Homo sapi
c 36	12	60.0	31	6	A43240	A43240 Scattered p
37	12	60.0	41	10	MM5TRNA	X-ray M musculus
c 38	12	60.0	45	6	AR099096	AR099096 Sequence
39	12	60.0	51	12	SYNEKRCK11	M-cell Synthetic
c 40	12	60.0	54	6	155061	155061 sequence 16
c 41	12	60.0	56	8	LIRRRE02	M82548 Trichodendro
c 42	12	60.0	60	6	E41553	E41553 Cecropin-de
c 43	12	60.0	60	9	BSEZPING03	AF054455 Homo sapi
c 44	11.8	59.0	20	6	AR084116	AR084116 Sequence
45	11.8	59.0	36	6	AX140180	AX140180 Sequence

ALIGNMENTS

RESULT 1
AX160629
LOCUS AX160629 51 bp DNA PAT 22-JUN-2001
DEFINITION Sequence 4957 from Patent WO0140521.
ACCESSION AX160629
VERSION AX160629.1 GI:14941960
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE 1 (bases 1 to 51)
AUTHORS Shimkets,R.A. and Leach,M.
TITLE Nucleic acids containing single nucleotide polymorphisms and methods of use thereof
JOURNAL Patent: WO 0140521 A 4957 07-JUN-2001;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source 1..51
/organism="Homo sapiens"

Contract/Grant No.: CO-56000; CO; NCI
Document type: Journal Article
Languages: ENGLISH
Main Citation Owner: NLM
Record type: Completed

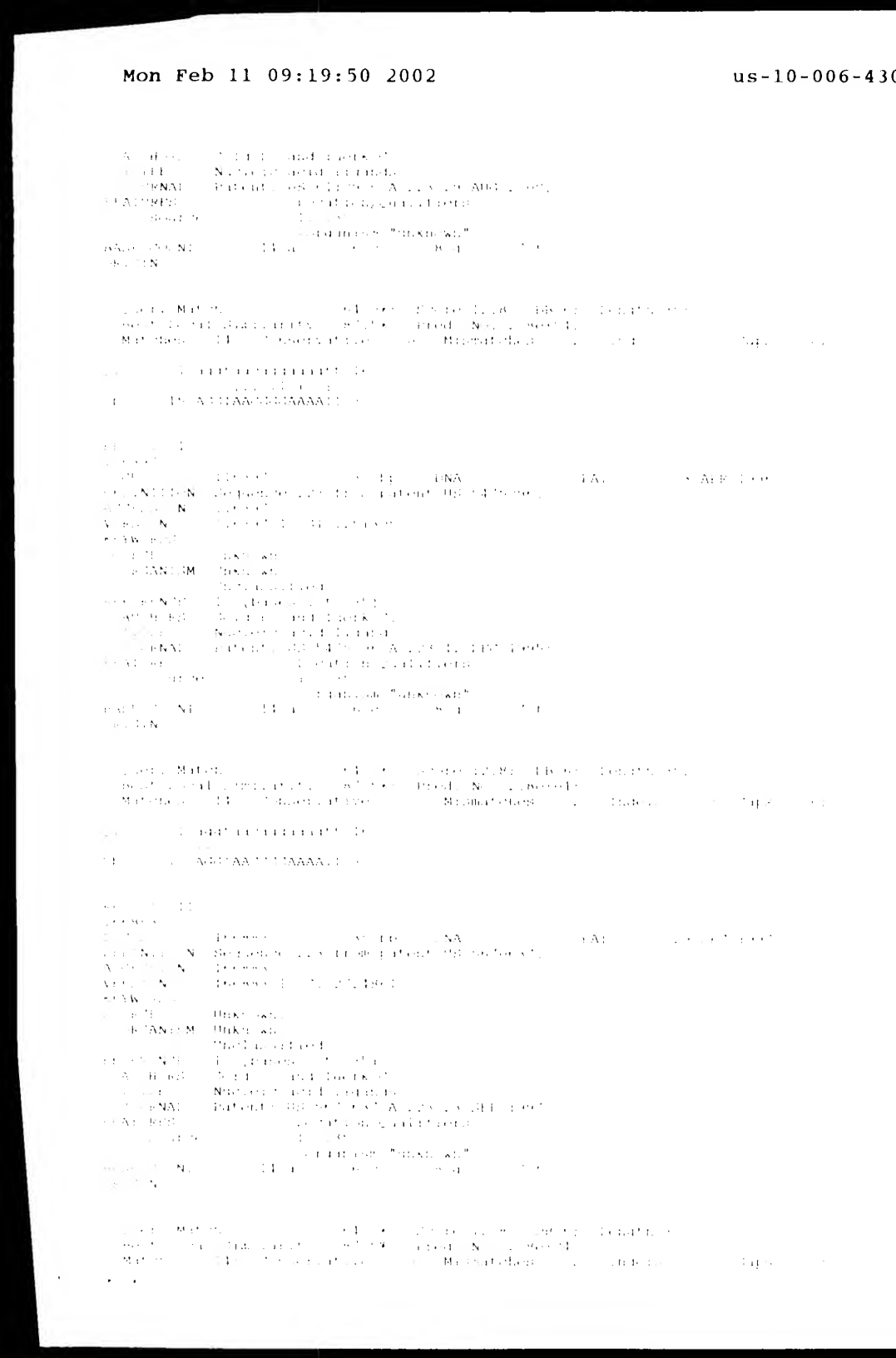
The envelope proteins of hepatitis C virus (HCV) are the likely targets of neutralizing **antibodies** and their molecular and functional characterization is relevant for vaccine development. We previously showed that surface-expressed E2 is a better immunogen than intracellular E2 and, therefore, we were interested in exploring more efficient ways to present E2 protein on the cell surface. We found that E2 targeted to the cell surface by replacement of its transmembrane domain did not bring E1 to the surface although E1 could be expressed independently on the cell surface if its transmembrane domain was similarly replaced. FACS analysis suggested that E2 expressed on the cell surface acquired its native conformation more efficiently when truncated at aa 661 than when truncated at aa 715. The shorter form of truncated E2 better retained the ability to bind the second extracellular loop (EC2) of CD81, the putative HCV receptor. Interestingly, deletion of the hypervariable region 1 (HVR1) did not perceptibly alter E2 structure; cell-surface forms of E2 lacking the HVR1 remained reactive with conformation-sensitive MAbs and were able to bind recombinant EC2 of CD81.

4/3,AB/20 (Item 20 from file: 155)
DIALOG(R)File 155:MEDLINE(R)

10823838 20363096 PMID: 10907850
Sequence-based structural features between Kvlqt1 and Tapa1 on mouse
on human 11p15.5: long-stretches of unusually
sequences of kvlqt1 between mouse and human.

Department of Biochemistry, Saga Medical

```
  Authors:  Chen,Y.Y., Hayashi,Y., Wu,I.J., Takaoka,E., Maekawa,K.,
            Watanabe,N., Tsuchiya,D., Hosoda,F., Arai,Y., Mizushima,H.,
            Motobashi,A., Ohira,M., Nakamoto,A., Liu,S., Hoshi,M., Horii,A.
            and Sonda,K.
  Title:    A bac based sts content map spanning a 35 mb region of human
            chromosome 1p35 p36
  Journal:  Genomics. 74 (1), 55-70 (2001)
  Medline:  21260102
REFERENCE   2  (bases 1 to 18)
  Authors:  Horii,A.
  Title:    Direct Submission
  Journal:  Submitted (04 AUG 2001) Akira Horii, Tohoku University School of
            Medicine, Molecular Pathology; 2-1 Seiryomachi, Aoba-ku, Sendai,
            Miyagi 980-8575, Japan (E-mail:horii@mail.cc.tohoku.ac.jp,
            Tel:81 22 717 8042, Fax:81 22 717 8047)
FEATURES             Location/Qualifiers
     source          1..18
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     misc_feature    1..18
                     /note="forward primer for human STS sts-cos51-17 at 1p36
                     sts-cos51-17 obtained from clones B4M23, B11G16, Human
                     BAC library RPCI-11"
BASE COUNT         5 a         3 c         7 g         3 t
ORIGIN


  Query Match             65.0%;  Score 13;  DB 12;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 2.1e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 ccgaaattcca 19
           |||||||||||
Db       4 CACGAAATTCAC 16


RESULT   6
AX060306
LOCUS       AX060306         19 bp    DNA             PAT       06-MAR-1998
DEFINITION  Sequence 1 from Patent WO9707240.
ACCESSION   AX060306
VERSION     AX060306.1  GI:3715281
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   Klatt,C. and Grobbee,D.E.
  TITLE     T1A POLYMORPHISM AND USE THEREOF IN DIAGNOSIS OF RISK OF THROMBUS
            ASSOCIATED DISEASE
  JOURNAL   Patent: WO 9707240 A 1 27-FEB-1997;
            AKZO NOBEL NV (NL)
FEATURES             Location/Qualifiers
     source          1..19
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT         5 a         5 c         4 g         5 t
ORIGIN


  Query Match             64.0%;  Score 12.8;  DB 6;  Length 19;
  Best Local Similarity   87.5%;  Pred. No. 2.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 atccaccgaaattcca 18
           |||||||| || |||||
Db       4 GTAACAGGACAGCTCA 19


RESULT   7
AX095468
LOCUS       AX095468         21 bp    DNA             PAT       08-MAR-2001
DEFINITION  Sequence 646 from Patent WO0118250.
ACCESSION   AX095468
VERSION     AX095468.1  GI:13511671
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Lander,E.S., Gargill,M., Ireland,J.S., Bolk,S., Daly,J., and
            Mccarthy,J.J.
  TITLE     Single nucleotide polymorphisms in genes
  JOURNAL   Patent: WO 0118250-A 646 15-MAR-2001;
            WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US) ; Millennium
            Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..21
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT         6 a         5 c         6 g         4 t         1 other.
ORIGIN


  Query Match             64.0%;  Score 12.8;  DB 6;  Length 21;
  Best Local Similarity   77.8%;  Pred. No. 2.7e+04;
  Matches   14;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 aqatcaccgaaattca 18
           |||| |:|||||| ||||
Db       4 AGCTGCCTGGCAAACTCA 21


RESULT   8
AR061541
LOCUS       AR061541         35 bp    DNA             PAT       24-SEP-1999
DEFINITION  Sequence 223 from patent US 5843653.
ACCESSION   AR061541
VERSION     AR061541.1  GI:5989232
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 35)
  AUTHORS   Gold,L. and Tuerk,C.
  TITLE     Method for detecting a target molecule in a sample by using a nucleic
            acid ligand
  JOURNAL   Patent: US 5843653 A 224 01-DEC-1998;
FEATURES             Location/Qualifiers
     source          1..35
                     /organism="unknown"
BASE COUNT         13 a        6 c         8 g         7 t
ORIGIN


  Query Match             64.0%;  Score 12.8;  DB 6;  Length 35;
  Best Local Similarity   87.5%;  Pred. No. 2.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 aqatcaccgaaaqtt 16
           |||||||| |||||.||
Db      15 AGGTAACAGGACAAAATT 30


RESULT   9
AR108440
LOCUS       AR108440         35 bp    DNA             PAT       14-FEB-2001
DEFINITION  Sequence 223 from patent US 6110900
ACCESSION   AR108440
VERSION     AR108440.1  GI:12824927
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 35)
```



```
RESULT 14
AX162440
LOCUS       AX162440                51 bp    DNA             PAT       22-JUN-2001
DEFINITION  Sequence 5758 from Patent WO0140521.
ACCESSION   AX162440
VERSION     AX162440.1  GI:14544761
KEYWORDS
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Lesch,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521 A 5758 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (5757 is other entry)
                     Accession number cq44039068"
BASE COUNT       20 a      2 c      11 g      18 t
ORIGIN


  Query Match             84.0%;  Score 12.6;  DB 6;  Length 51;
  Best Local Similarity   87.5%;  Pred. No. 2.8e-04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 cagtaacagtaaagtt 16
           ||| || ||||||||
Db      28 AGGAAAGAGGAAAGTT 43


RESULT 15
AR121249
LOCUS       AR121249                51 bp    DNA             PAT       16-MAY-2001
DEFINITION  Sequence 29 from patent US 6159710.
ACCESSION   AR121249
VERSION     AR121249.1  GI:14104825
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Fraser,N.W., Zabolotny,J.M. and Krummenacher,C.F.
  TITLE     Method and Compositions for stabilizing unstable gene transcripts
  JOURNAL   Patent: US 6159710 A 29-12-DEC-2000;
FEATURES             Location/Qualifiers
     source          1..51
                     /organism="unknown"
BASE COUNT       9 a      14 c       4 g      10 t
ORIGIN


  Query Match             83.0%;  Score 12.6;  DB 6;  Length 51;
  Best Local Similarity   78.9%;  Pred. No. 4.5e-04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 cagtaacagtaaagttcag 19
           ||  || ||||||| ||||
Db      21 AGACAAGAGGAAAGGTCAG 3


Search completed: February 5, 2002, 01:44:02
Job time : 21.01 sec
```

SeqWare version 4.5

Run mode: nucleic search, using sw model

Run on: February 9, 2002, 01:44:02 ; Search time 2519.39 Seconds
(without alignments)
130.962 Million cell updates/sec

Title: US-10-006-430-61
Perfect score: 29
Sequence: 1 catcagcctgaagtttca 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 147240 seqs, 82405e+256 residues

Total number of hits satisfying chosen parameters: 986446

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_oz:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_oz:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgs_hum:*
31: em_htgs_inv:*
32: em_htgs_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	14	70.4	27	6	AR040450	AR040450 Sequence
2	14.8	69.0	89	6	E38610	E38610 Method for
c 3	13	65.0	46	6	AR045322	AR045... Sequence
c 4	13	65.0	46	6	AR141808	AR141808 Sequence
c 5	13	65.0	46	6	152474	152474 Sequence 11
c 6	13	65.0	46	6	161995	161995 Sequence 54
c 7	13	65.0	48	6	AR046709	AR046709 Sequence
c 8	13	65.0	48	6	AR047647	AR047647 Sequence
c 9	13	65.0	48	6	AR047799	AR047799 Sequence
c 10	13	65.0	48	6	154851	154851 Sequence 19
c 11	13	65.0	48	6	154699	154699 Sequence 24
c 12	13	65.0	48	6	154851	154851 Sequence 2?
c 13	12.8	64.0	45	6	A05540	A05540 collagen gene
14	12.8	64.0	45	6	A05548	A05548 cell surface
c 15	12.8	64.0	45	6	AR001633	AR001633 Sequence
16	12.8	64.0	45	6	AR001641	AR001641 Sequence
c 17	12.8	64.0	45	6	109090	109090 Sequence 5
18	12.8	64.0	45	6	109098	109098 Sequence 13
19	12.6	63.0	48	6	AX085617	AX085617 Sequence
c 20	12.6	63.0	57	6	AR086612	AR086612 Sequence
21	12.4	62.0	19	6	122556	122556 Sequence 12
22	12.4	62.0	19	6	147481	147481 Sequence 14
c 23	12.4	62.0	22	6	AR026879	AR026879 Sequence
c 24	12.4	62.0	22	6	AR049405	AR049405 Sequence
c 25	12.4	62.0	22	6	AR065563	AR065563 Sequence
26	12.4	62.0	22	6	AX149450	AX149450 Sequence
c 27	12.4	62.0	27	6	AR040452	AR040452 Sequence
c 28	12.4	62.0	38	6	AR057878	AR057878 Sequence
c 29	12.4	62.0	38	6	AR058151	AR058151 Sequence
c 30	12.4	62.0	38	6	AR115636	AR115636 Sequence
c 31	12.4	62.0	38	6	AR115909	AR115909 Sequence
32	12.4	62.0	49	6	122755	122755 Sequence 24
33	12.4	62.0	49	6	147580	147580 Sequence 24
34	12.4	62.0	45	6	147719	147719 Sequence 18
c 35	12.4	62.0	57	9	BSU14021	U14021 Bacillus [rec]
c 36	12.2	61.0	22	4	D87944602	D87944602 Oryctolag
c 37	12.2	61.0	30	6	AX1711	AX1711 Mus musculus
c 38	12.2	61.0	30	6	AR059101	AR059101 Sequence
c 39	12.2	61.0	30	6	120690	120690 Sequence 5
40	12.2	61.0	38	6	147782	147782 Sequence 29
41	12.2	61.0	38	6	194642	194642 Sequence 29
c 42	12.2	61.0	40	6	AR059102	AR059102 Sequence
c 43	12.2	61.0	40	6	AR059103	AR059103 Sequence
c 44	12.2	61.0	40	6	AR059106	AR059106 Sequence
c 45	12.2	61.0	60	6	AX147742	AX147742 Sequence

ALIGNMENTS

RESULT 1
AR040450/c
LOCUS AR040450 27 bp DNA PAT 22-SEP-1999
DEFINITION Sequence 1298 from patent US 5807744.
ACCESSION AR040450
VERSION AR040450.1 GI:5959813
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 27)
AUTHORS Stinchcomb,D.T. and McSwiggen,J.A.
TITLE Interleukin 2 receptor gamma chain ribozyme
JOURNAL Patent: US 5807744 A 1298 15-SEP-1998;
FEATURES Location/Qualifiers
source 1..27
/organism="unknown"
BASE COUNT 9 a 6 c 7 g 5 t 0 others
ORIGIN

```
RESULT   6
L54851/c
LOCUS       L54851          38 bp    DNA             PAT       07-OCT-1997
DEFINITION  Sequence 549 from patent US 5658780.
ACCESSION   L54851
VERSION     L54851.1  GI:2477043
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Stinchcomb,D.T., Draper,K.Y. and McSwiggen,J.
  TITLE     Rel 1 Targeted ribozymes
  JOURNAL   Patent: US 5658780 A 549 19-AUG-1997;
FEATURES             Location/Qualifiers
     source          1..38
                     /organism="unknown"
BASE COUNT       10 a       9 c      12 g       7 t
ORIGIN


  Query Match             65.0%;  Score 13;  DB 6;  Length 38;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 catcagccctgaa 13
           |||||||||||||
Db      14 CATCAGCCCTGAA 1


RESULT   7
AR047799/c
LOCUS       AR047799        38 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 1592 from patent US 5817796.
ACCESSION   AR047799
VERSION     AR047799.1  GI:5968264
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Stinchcomb,D.T., Draper,K., McSwiggen,J. and Jarvis,E.
  TITLE     C-myb ribozymes having 2' 5' linked adenylate residues
  JOURNAL   Patent: US 5817796 A 1592 06-OCT-1998;
FEATURES             Location/Qualifiers
     source          1..38
                     /organism="unknown"
BASE COUNT       10 a       9 c      12 g       7 t
ORIGIN


  Query Match             65.0%;  Score 13;  DB 6;  Length 38;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 catcagccctgaa 13
           |||||||||||||
Db      14 CATCAGCCCTGAA 2


RESULT   8
AR047647/c
LOCUS       AR047647        38 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 2440 from patent US 5817796.
ACCESSION   AR047647
VERSION     AR047647.1  GI:5969112
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Stinchcomb,D.T., Draper,K., McSwiggen,J. and Jarvis,E.
  TITLE     C-myb ribozymes having 2' 5' linked adenylate residues
  JOURNAL   Patent: US 5817796 A 2440 06-OCT-1998;
FEATURES             Location/Qualifiers
     source          1..38
                     /organism="unknown"
BASE COUNT       10 a       9 c      12 g       7 t
ORIGIN


  Query Match             65.0%;  Score 13;  DB 6;  Length 38;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 catcagccctgaa 13
           |||||||||||||
Db      14 CATCAGCCCTGAA 2


RESULT   9
AR047799/c
LOCUS       AR047799        38 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 2592 from patent US 5817796.
ACCESSION   AR047799
VERSION     AR047799.1  GI:5969264
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Stinchcomb,D.T., Draper,K., McSwiggen,J. and Jarvis,E.
  TITLE     C-myb ribozymes having 2' 5' linked adenylate residues
  JOURNAL   Patent: US 5817796 A 2592 06-OCT-1998;
FEATURES             Location/Qualifiers
     source          1..38
                     /organism="unknown"
BASE COUNT       10 a       9 c      12 g       7 t
ORIGIN


  Query Match             65.0%;  Score 13;  DB 6;  Length 38;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 catcagccctgaa 13
           |||||||||||||
Db      14 CATCAGCCCTGAA 2


RESULT  10
L54851/c
LOCUS       L54851          38 bp    DNA             PAT       07-OCT-1997
DEFINITION  Sequence 1592 from patent US 5649042.
ACCESSION   L54851
VERSION     L54851.1  GI:2477054
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Stinchcomb,D.T., Draper,K., McSwiggen,J. and Jarvis,E.
  TITLE     C-myb targeted ribozymes
  JOURNAL   Patent: US 5649042-A 1592 08-JUL-1997;
FEATURES             Location/Qualifiers
     source          1..38
                     /organism="unknown"
BASE COUNT       10 a       9 c      12 g       7 t
ORIGIN


  Query Match             65.0%;  Score 13;  DB 6;  Length 38;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

/organism "unknown"
QY 1 AGNT 9 9 8 19
RESIN

Query Match 64.0%; Score 12.6; DB 6; Length 45;
Best Local Similarity 87.5%; Pred. No. 3.4e-04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 [illegible] 18
| | | | | | | | | | | | | | |
Db 68 [illegible] 23

Search completed: February 9, 2002, 01:44:03
Job time: 21902 sec

GeneCore version 4.5

GM nucleic - nucleic search, using sw model

Run on February 6, 2002, 01:44:03 ; Search time 2519.39 Seconds
(without alignments)
180.962 Million cell updates/sec

Title: US-10-006-430-63
Perfect score: 20
Sequence: 1 gaaaggaaaaaaaaaaaaaa 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 824850775 residues

Total number of hits satisfying chosen parameters: 586436

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_ov:*
21: em_ov2:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

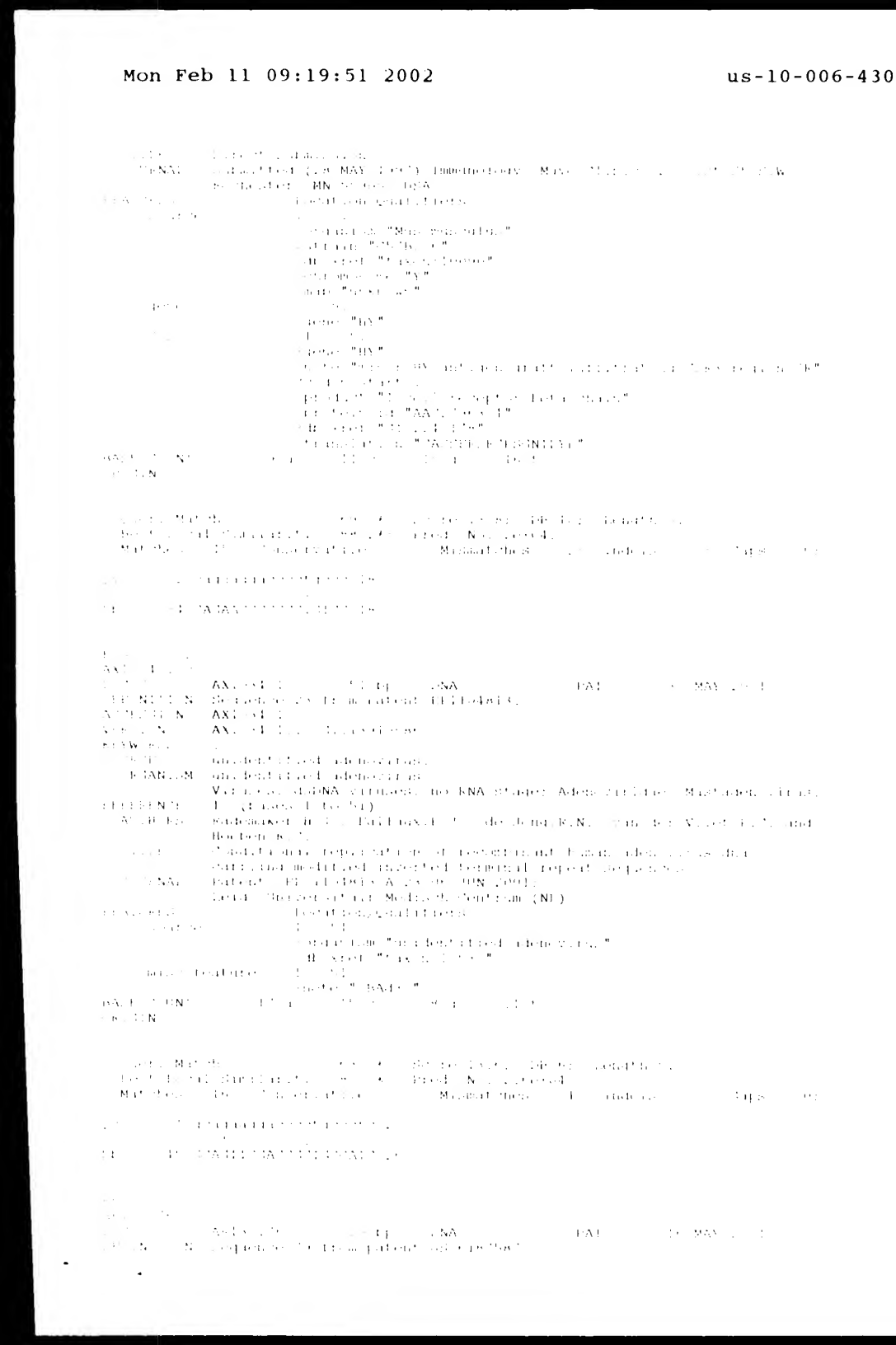
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	13.8	69.0	51	10	AF005618	AF005618 Mus muscu
c 2	13.6	68.0	51	6	AX103401	AX103401 Sequence
3	13.4	67.0	20	6	AR130176	AR130176 Sequence
c 4	13.4	67	47	6	AR081411	AR081411 Sequence
5	13	65.0	23	6	176404	176404 Sequence 19
6	13	65.0	23	6	183824	183824 Sequence 19
7	13	65.0	23	6	186152	186152 Sequence 19
8	13	65.0	23	6	186246	186246 Sequence 19
c 9	12.8	64.0	22	6	A39978	A39978 Sequence 11
c 10	12.8	64.0	22	6	123504	123504 Sequence 11
c 11	12.8	64.0	29	6	E26547	E26547 DiffSI seq
c 12	12.8	64.0	41	6	112942	112942 Sequence 2
13	12.8	64.0	41	6	112943	112943 Sequence 2
c 14	12.8	64.0	45	22	E11374	E11374 PCR primer
c 15	12.8	64.0	48	22	E09865	E09865 Primer 9/2
c 16	12.8	64.0	51	6	AX160530	AX160530 Sequence
17	12.6	63.0	22	6	E15011	E15011 PCR primer
18	12.6	63.0	22	6	165401	165401 Sequence 2
c 19	12.6	63.0	48	6	E17013	E17013 DNA encodin
c 20	12.6	63.0	50	6	AR032797	AR032797 Sequence
c 21	12.6	63.0	50	6	129537	129537 Sequence 40
c 22	12.6	63.0	50	6	191211	191211 Sequence 40
c 23	12.6	63.0	51	6	AX115281	AX115281 Sequence
24	12.4	62.0	20	6	E29938	E29938 HIV infecte
25	12.4	62.0	29	6	AX118710	AX118710 Sequence
c 26	12.4	62.0	50	6	AX118707	AX118707 Sequence
27	12.2	61.0	34	6	AR119727	AR119727 Sequence
28	12.2	61.0	34	6	AR129857	AR129857 Sequence
c 29	12.2	61.0	38	6	AX000273	AX000273 Sequence
30	12.2	61.0	44	6	AR086647	AR086647 Sequence
c 31	12.2	61.0	48	9	HSBAFIB	X71495 H.Sapiens (
32	12.2	61.0	50	6	AR042514	AR042514 Sequence
33	12.2	61.0	50	6	129254	129254 Sequence 12
34	12.2	61.0	50	6	190928	190928 Sequence 12
c 35	12.2	61.0	51	6	AX156931	AX156931 Sequence
c 36	12.2	61.0	51	6	AX156932	AX156932 Sequence
c 37	12.2	61.0	51	9	HSA224215	A1224215 Homo Sapi
38	12	60.0	15	6	176403	176403 Sequence 18
39	12	60.0	15	6	183823	183823 Sequence 18
40	12	60.0	15	6	186151	186151 Sequence 18
41	12	60.0	15	6	186245	186245 Sequence 18
42	12	60.0	19	6	176391	176391 Sequence 6
43	12	60.0	19	6	183811	183811 Sequence 6
44	12	60.0	19	6	186139	186139 Sequence 6
45	12	60.0	19	6	186233	186233 Sequence 6

ALIGNMENTS

RESULT 1
AF005618/c
LOCUS AF005618 51 bp mRNA ROD 31-JUL-1998
DEFINITION Mus musculus T cell receptor beta chain, minor BY antigen-spati
infiltrating CDR4 region 2B, mRNA, partial cds.
ACCESSION AF005618
VERSION AF005618.1 GI:2240177
KEYWORDS .
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 51)
AUTHORS Johnston,S.L. and Wettstein,P.J.
TITLE Spectratyping of TCR expressed by CTL infiltrating male antigen (HY) disparate allografts
JOURNAL J. Immunol. 160 (7), 3474-3484 (1998)
MEDLINE 98189784
REFERENCE 2 (bases 1 to 51)
AUTHORS Johnston,S.L., Borson,N. and Wettstein,P.J.



BASE COUNT 4 a 11 c 5 g 3 t
ORIGIN

Query Match 65.0%; Score 13; DB 6; Length 23;
Best Local Similarity 100.0%; Pred. No. 6e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 gaaccccctgcccct 19
||||||||||||||
Db 11 GGACCCCCTGCCCCT 23

RESULT 6
186246
LOCUS 186246 23 bp DNA PAT 10-AUG-1998
DEFINITION Sequence 19 from patent US 5714595.
ACCESSION 186246
VERSION 186246.1 GI:3205964
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 23)
AUTHORS Mak,P. and Karathanasis,S.K.
TITLE Mechanism based screen for retinoid X receptor agonists and antagonists
JOURNAL Patent: US 5714595 A 19-FEB-1998;
FEATURES location/qualifiers
source 1..23
/organism="unknown"
BASE COUNT 4 a 11 c 5 g 3 t
ORIGIN

Query Match 65.0%; Score 13; DB 6; Length 23;
Best Local Similarity 100.0%; Pred. No. 6e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 gaaccccctgcccct 19
||||||||||||||
Db 11 GGACCCCCTGCCCCT 23

RESULT 7
186152
LOCUS 186152 23 bp DNA PAT 10-JUN-1998
DEFINITION Sequence 19 from patent US 5760650.
ACCESSION 186152
VERSION 186152.1 GI:3205824
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 23)
AUTHORS Mak,P. and Karathanasis,S.K.
TITLE Mechanism based screen for retinoid X receptor agonists and antagonists
JOURNAL Patent: US 5760650 A 19-JUN-1998;
FEATURES location/qualifiers
source 1..23
/organism="unknown"
BASE COUNT 4 a 11 c 5 g 3 t
ORIGIN

Query Match 65.0%; Score 13; DB 6; Length 23;
Best Local Similarity 100.0%; Pred. No. 6e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 gaaccccctgcccct 19
||||||||||||||
Db 11 GGACCCCCTGCCCCT 23

RESULT 8
186246
LOCUS 186246 23 bp DNA PAT 10-JUN-1998
DEFINITION Sequence 19 from patent US 5760682.
ACCESSION 186246
VERSION 186246.1 GI:3205964
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 23)
AUTHORS Mak,P. and Karathanasis,S.K.
TITLE Mechanism based screen for retinoid X receptor agonists and antagonists
JOURNAL Patent: US 5760682 A 19-JUN-1998;
FEATURES location/qualifiers
source 1..23
/organism="unknown"
BASE COUNT 4 a 11 c 5 g 3 t
ORIGIN

Query Match 65.0%; Score 13; DB 6; Length 23;
Best Local Similarity 100.0%; Pred. No. 6e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 gaaccccctgcccct 19
||||||||||||||
Db 11 GGACCCCCTGCCCCT 23

RESULT 9
A39978/c
LOCUS A39978 22 bp DNA PAT 5-MAR-1997
DEFINITION Sequence 11 from Patent WO9421790.
ACCESSION A39978
VERSION A39978.1 GI:2250173
KEYWORDS .
SOURCE unidentified.
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 22)
AUTHORS Hayden,M., Goldberg,P., Andrew,S., Rommens,J.M., Tsui,L.-C. and Buchwald,M.
TITLE Process for isolating genes and the gene causative of a heritable disease and differential 5' polyadenylation in the gene
JOURNAL Patent: WO 9421790 A 11-26 SEP 1994;
UNIV BRITISH COLUMBIA (CA)
COMMENT other publication CA 2092455 940926
other publication AU 6281094 941011
other publication US 5534438 960709
other publication CA 2158918 940929.
FEATURES location/qualifiers
source 1..22
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 4 a 6 c 6 g 6 t
ORIGIN

Query Match 64.0%; Score 12.8; DB 6; Length 22;
Best Local Similarity 87.5%; Pred. No. 7.7e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 4 gaagcccctgcccct 19
||| || |||||||||
Db 18 GAAIGAGCCCTGCCCT 3

RESULT 10
123504/c
LOCUS 123504 22 bp DNA PAT 27-OCT-1995

```
ID   receptor.
XX
AC   JP 1996131172 A/4.
XX
DE   unidentified
CC   unclassified.
XX
RN   [1]
RP   1,45
RA   ... Y., Anaguchi H.;
RT   "DNA CODING PROTEIN BC IN LIGAND BOUND REGION IN GRANULOCYTE COLONY
RT   STIMULATING FACTOR RECEPTOR";
RL   Patent number JP1996131172 A/4, 28 MAY 1996,
RL   FANPAKU KAGAKU KENKYUSHO**P.
XX
CC   OS   None
CC   OC   Artificial sequences.
CC   PN   JP 1996131172 A/4
CC   PD   28 MAY 1996
CC   PF   14 NOV 1994 JP 1994278041
CC   PI   OTA YOSHIMI, ANAGUCHI HIROYUKI
CC   DR   C12N15/09,C12P21/02,C12R1:91),(C12P21/02,C12R1:19);
CC   CC   strandedness: Single;
CC   CC   topology: Linear;
CC   CC   hypothetical: No;
CC   FH   Key                 Location/qualifiers
CC   FH
CC   FT   source              1..48
CC   FT                       /organism "Artificial sequences"
XX
FH   Key               Location/qualifiers
FH
FT   source            1..48
FT                     /db_xref "taxon:2644"
FT                     /organism "unidentified"
XX
SQ   Sequence 48 BP; 10 A; 16 C; 14 G; 9 T; 0 other;


  Query Match             64.0%;  Score 12.8;  DB 22;  Length 48;
  Best Local Similarity   87.5%;  Prod. No. 6.9e-04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 aaaaccccctgcccc 20
           ||| |||||| |||||
Db      27 AAGAACCCCTACCCC 12


Search completed: February  9, 2002, 01:44:04
Job time: 21893 sec
```

```
ID   receptor.
XX
AC   JP 1996131172 A/4.
XX
DE   unidentified
CC   unclassified.
XX
RN   [1]
RP   1,45
RA   ... Y., Anaguchi H.;
RT   "DNA CODING PROTEIN BC IN LIGAND BOUND REGION IN GRANULOCYTE COLONY
RT   STIMULATING FACTOR RECEPTOR";
RL   Patent number JP1996131172 A/4, 28 MAY 1996,
RL   FANPAKU KAGAKU KENKYUSHO**P.
XX
CC   OS   None
CC   OC   Artificial sequences.
CC   PN   JP 1996131172 A/4
CC   PD   28 MAY 1996
CC   PF   14 NOV 1994 JP 1994278041
CC   PI   OTA YOSHIMI, ANAGUCHI HIROYUKI
CC   DR   C12N15/09,C12P21/02,C12R1:91),(C12P21/02,C12R1:19);
CC   CC   strandedness: Single;
CC   CC   topology: Linear;
CC   CC   hypothetical: No;
CC   FH   Key                 Location/qualifiers
CC   FH
CC   FT   source              1..48
CC   FT                       /organism "Artificial sequences"
XX
FH   Key               Location/qualifiers
FH
FT   source            1..48
FT                     /db_xref "taxon:2644"
FT                     /organism "unidentified"
XX
SQ   Sequence 48 BP; 10 A; 16 C; 14 G; 9 T; 0 other;


  Query Match             64.0%;  Score 12.8;  DB 22;  Length 48;
  Best Local Similarity   87.5%;  Prod. No. 6.9e-04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 aaaaccccctgcccc 20
           ||| |||||| |||||
Db      27 AAGAACCCCTACCCC 12


Search completed: February  9, 2002, 01:44:04
Job time: 21893 sec
```

```
ID   1P
Locus/
AC   1P0065;
XX
AC   1P0065;
XX
SV   1P0065.1
XX
DT   08-OCT-1997 (Rel. 52, Created)
DT   07-SEP-2000 (Rel. 65, Last updated, Version 2)
XX
DE   Primer.
XX
KW   JP 1995227288 A/10.
XX
OS   unidentified
CC   unclassified.
XX
RN   [1]
RP   1-18
RA   OTA Y., Anaguchi H., Hiraoka ...
RT   "DNA CODING FOR LIGAND BOND RANGE PROTEIN OF GRANULOCYTE COLONY STIMULATION
```

```
RT   FACTOR RECEPTOR";
RL   Patent number JP1995227288 A/10, 29 AUG 1995,
RL   FANPAKU KOGAKU KENKYUSHO:FP.
XX
CC   OS   None
CC   OC   Artificial sequences.
CC   PN   JP 1995227288 A/10
CC   PD   29 AUG 1995
CC   PF   30 MAY 1994 JP 1994116252
CC   PR   21 DEC 1993 JP 93P   321862
CC   PI   OTA YOSHIMI, ANAGUCHI HIROYUKI, HIRAOKA OSAMU
CC   DR   C12N15/09,C12P21/02,(C12N15/09,C12R1:91),(C12P21/02,C12R1:19);
CC   CC   strandedness: Single;
CC   CC   topology: Linear;
CC   CC   hypothetical: No;
CC   FH   Key                 Location/qualifiers
CC   FH
CC   FT   source              1..48
CC   FT                       /organism "Artificial sequences"
XX
FH   Key               Location/qualifiers
FH
FT   source            1..48
FT                     /db_xref "taxon:2644"
FT                     /organism "unidentified"
XX
SQ   Sequence 48 BP; 10 A; 16 C; 14 G; 9 T; 0 other;


  Query Match             64.0%;  Score 12.8;  DB 22;  Length 48;
  Best Local Similarity   87.5%;  Prod. No. 6.9e-04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 aaaaccccctgcccc 20
           ||| |||||| |||||
Db      27 AAGAACCCCTACCCC 12


Search completed: February  9, 2002, 01:44:04
Job time: 21893 sec
```

Run on: February 9, 2002, 01:01:51 ; Search time 8488.86 Seconds
(without alignments)
25,619 Million cell updates/sec

Title: US-10-006-430-68
Perfect score: 60
Sequence: 1 atgcaatccaaaaaaaaaaa 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 11551937 seqs, 5472864281 residues

Total number of hits satisfying chosen parameters: 111874

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estome:*
5: em_estpln:*
6: em_estbac:*
7: em_estvrs:*
8: em_estoay:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_plt:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Prob. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

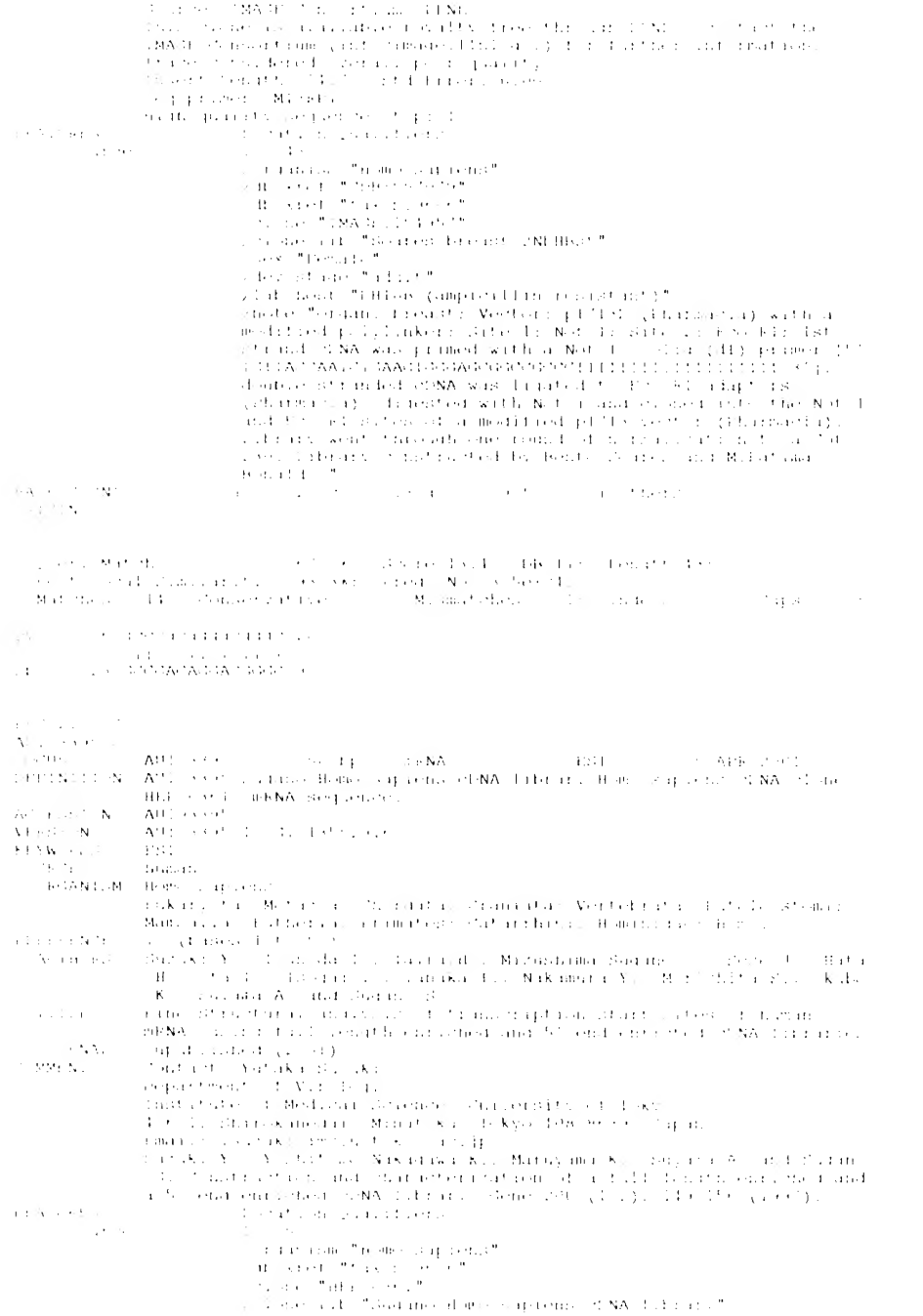
SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	15.4	77.0	60	10	AI905446	AI905446 CM_BI086
2	14.2	71.0	50	10	AU106349	AU106349 AU106349
3	13.6	68.0	60	13	AI465424	AI465424 E. bruneci
4	13.4	67.0	47	11	W05633	W05633 za90d08.r1
5	13.4	67.0	48	10	AI174097	AI174097 yz87e06.r
6	13.4	67.0	48	11	R55448	R55448 yj79d07.r1
7	13.2	66.0	50	10	AU103395	AU103395 AU103395
8	13	65.0	42	13	AZ941495	AZ941495 ZM020IF16
9	12.8	64.0	50	11	F16075	F16075 yb98102.r1
10	12.6	63.0	42	13	AL488799	AL488799 T. brucei
11	12.6	63.0	44	13	AZ492829	AZ492829 ZM032F01
12	12.2	61.0	41	11	H45796	H45796 HUMGS03018
13	12.2	61.0	52	10	AI214648	AI214648 qa52a11.x
14	12.2	61.0	52	10	AI439245	AI439245 tf72c05.x
15	12.2	61.0	54	11	BG046071	BG046071 ho75e04
16	12.2	61.0	55	10	AA948457	AA948457 ne59e08.s
17	12	60.0	40	10	AI152480	AI152480 od85f01.r
18	12	60.0	43	10	AI423946	AI423946 tf15e05.x
19	12	60.0	50	11	BE149814	BE149814 601488B02
20	12	60.0	53	13	AZ825842	AZ825842 ZM_IcHE1
21	12	60.0	55	13	AZ634522	AZ634522 ZM340E21
22	12	60.0	57	13	AZ472075	AZ472075 ZM0282A08
23	11.8	59.0	32	13	AZ769468	AZ769468 ZM0
24	11.8	59.0	36	13	AZ828771	AZ828771 ZM1
25	11.8	59.0	37	10	AA458001	AA458001 zt74d10.r
26	11.8	59.0	46	10	AA244435	AA244435 nc61b09.s
27	11.8	59.0	49	10	AA657082	AA657082 zt24d09.r
28	11.8	59.0	50	10	AU106309	AU106309 AU106309
29	11.8	59.0	50	10	AU107638	AU107638 AU107638
30	11.8	59.0	51	11	BG913750	BG913750 he61e04
31	11.8	59.0	51	13	AZ433814	AZ433814 ZM_19F07
32	11.8	59.0	55	10	AI283325	AI283325 qt14d11.x
33	11.8	59.0	55	13	IAI41D03	AI465424 E. bruneci
34	11.8	59.0	57	11	T68090	T68090 yc54c05.sl
35	11.6	58.0	41	10	AI394804	AI394804 tf67d12.x
36	11.6	58.0	44	10	AI584193	AI584193 tf84c03.x
37	11.6	58.0	49	10	BE544028	BE544028 601
38	11.6	58.0	46	10	AI284114	AI284114 qt 2e11.x
39	11.6	58.0	47	13	IAI95A08	AI477911 E. bruneci
40	11.6	58.0	49	10	AA677795	AA677795 zt24d06.s
41	11.6	58.0	49	10	AA781646	AA781646 ai50a11.s
42	11.6	58.0	50	10	AU102582	AU102582 AU102582
43	11.6	58.0	50	10	AU104781	AU104781 AU104781
44	11.6	58.0	50	11	BE571906	BE571906 601
45	11.4	57.0	26	13	AZ775254	AZ775254 ZM0007B11

ALIGNMENTS

RESULT 1
AI905446
LOCUS AI905446 60 bp mRNA EST 09-MAR-2000
DEFINITION CM_BI086 170299.085 BI086 Homo sapiens cDNA, mRNA sequence
ACCESSION AI905446
VERSION AI905446.1 GI:6495733
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE 1 (bases 1 to 60)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,A. M., Zago,M.A., Bordin,S., Costa Cruz,I.,
Coldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.R.,
Bromstein,A., deOliveira,P.S., Pircher,P., Jones,T.C., V Heardon,
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo SP,
Brazil
Tel: +55 11 2704922
Fax: +55 11 2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project, this entry can be seen in the following URL
(http://www.ludwig.org.br/seq/estlist.phtml?CM&2_CM_BI086_08.html
&3_170299&14_1)
Seq primer: puc 18 forward.

```
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, Mo 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty free through LENL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: mob.PECA-F1
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1..43
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /db_xref="clone:IMAGE:29023"
                     /clone="IMAGE:29023"
                     /clone_lib="Soares fetal lung NbHL19W"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer
                     [5' TGTTAGCAAACTCGAAGTGGGAGCGGCCGCAATTTTTTTTTTTTTTTT 3'],
                     double stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. This library was constructed
                     from the same fetus as the fetal heart library, Soares
                     fetal heart NbHHT9W."
BASE COUNT          14 a      15 c      15 g       5 t
ORIGIN


  Query Match             67.0%;  Score 14.4;  DB 11;  Length 43;
  Best Local Similarity   93.3%;  Prob. No. 3.5e-04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 tgtgcagcccacagga 19
           ||||||||| ||||
Db      12 GTGCAGCCCTCAGGA 26


RESULT  5
AI174097
LOCUS       AI174097          443 bp    mRNA            EST       07-OCT-1998
DEFINITION  vz87c06.r1 Soares mammary gland_NbMMG Mus musculus cDNA clone
            IMAGE:1433450 5' similar to TR:O35459 O35459 ECHIP. ;, mRNA
            sequence.
ACCESSION   AI174097
VERSION     AI174097.1  GI:4720260
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Shopkar,M., Tan,F., Underwood,K., Moore,B.,
            Theisen,B., Wylie,L., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The Wash-U RIKEN Mouse EST Project
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Marra M/Mouse EST Project
            Wash-U RIKEN Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, Mo 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty free through LENL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:692994
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: 28mlk rev2 KS from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1..43
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1433450"
                     /clone_lib="Soares mammary gland NbMMG"
                     /sex="male"
                     /tissue_type="mammary gland"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Organ: mammary gland; Vector: pT7T3-Pac (Pharmacia
                     ) with a modified polylinker; Site 1: Not I; Site 2: Eco
                     RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                     primer [5'
                     TGTTAGCAAACTCGAAGTGGGAGCGGCCGCAATTTTTTTTTTTTTTTT
                     3']; double stranded cDNA was ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     RNA provided by Dr. Minori Ko, Wayne State Univ., Library
                     constructed and normalized by Bento Soares and M Fatima
                     Bonaldo."
BASE COUNT          8 a      14 c      14 g       7 t
ORIGIN


  Query Match             67.0%;  Score 14.4;  DB 10;  Length 43;
  Best Local Similarity   93.3%;  Prob. No. 3.5e-04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 atgcagcccacagga 15
           ||||||||| ||||
Db      14 GTGCAGCCCTCAGGA 28


RESULT  6
X55440/c
LOCUS       X55440            443 bp    mRNA            EST       22-MAY-1995
DEFINITION  yj79d07.r1 Soares breast 2NbHBst Homo sapiens cDNA clone
            IMAGE:154927 5' similar to gb:K01228 PROCOLLAGEN ALPHA 1(I) CHAIN
            PRECURSOR (HUMAN);contains MER22 repetitive element ;, mRNA
            sequence.
ACCESSION   X55440
VERSION     X55440.1  GI:824735
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The Wash-U-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, Mo 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1417
            High quality sequence start: 1
            High quality sequence stop: 1
```



and selected for ampicillin resistance."
BASE COUNT 4 a 8 c 14 g 16 t
ORIGIN

Query Match 65.0%; Score 13; DB 13; Length 42;
Best Local Similarity 100.0%; Pred. No. 5.3e-04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 cccccacaaccacc 16
||||||||||||| ||
Db 16 CACCCACACGCAC 6

RESULT 9
AI630975/c
LOCUS AI630975 50 bp mRNA EST 16-FEB-1999
DEFINITION yb80107.r1 Stratagene lung (#937210) Homo sapiens cDNA clone IMAGE:79247 5' similar to gb:Kot144 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, GAMMA CHAIN PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION AI630975
VERSION AI630975.1 GI:6667342
KEYWORDS EST.
SEGMENT human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Hillier L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore ,B., Morris,M., Parsons,J., Prange,C., Ritkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry Meg,J., Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, Mo 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 849
High quality sequence starts: 1 High quality sequence stops: 1
Source: IMAGE Consortium, LLNL This clone is available royalty free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Trace considered overall poor quality
Insert Length: 849 Std Error: 0.00
Seq primer: M13RPI
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1..50
/organism="Homo sapiens"
/db_xref="GDB:482854"
/db_xref="taxon:9606"
/clone="IMAGE:79247"
/clone_lib="Stratagene lung (#937210)"
/sex="male"
/dev_stage="72 years"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI ; Site_2: XhoI; Cloned unidirectionally. Primer: oligo dT, normal lung. Average insert size: 1.0 kb; Uni ZAP XR Vector; 5' adaptor sequence: 5' GAATTCGGCACGAG 3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
BASE COUNT 13 a 12 c 12 g 10 t 1 others
ORIGIN

Query Match 64.0%; Score 12.8; DB 11; Length 50;
Best Local Similarity 87.5%; Pred. No. 6.5e-04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 cccccacaaccacc 17
||||||| | ||||||
Db 42 CACCCACGAACACCGG 27

RESULT 10
TA246E04Q/c
LOCUS TA246E04Q 32 bp DNA GSS 13-OCT-2000
DEFINITION T. brucei sheared genomic DNA clone 246e04, reverse sequence, genomic survey sequence.
ACCESSION AL488799
VERSION AL488799.1 GI:11864905
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
REFERENCE 1 (bases 1 to 32)
AUTHORS Hall,N., Berriman,S., Lennard,N.J., Ivens,A., Aert,R., Chillingworth,C., Ormond,D., Harris,D., El-Sayed,N., Ghedin ,E., Melville,S.E., Rajandream,M.A. and Barrell,B.G.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, England (E-mail: barrell@sanger.ac.uk and nh1@sanger.ac.uk)
COMMENT Constructed at the Institute for Genomic Research (TIGR) Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The vector method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vandin and B. Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES Location/Qualifiers
source 1..32
/organism="Trypanosoma brucei"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="246e04"
BASE COUNT 3 a 16 c 4 g 10 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 13; Length 32;
Best Local Similarity 78.9%; Pred. No. 7.8e-04;
Matches 15; Conservative 0; Mismatches 4; Indels 2; Gaps 0;

Qy 1 qtgcaccccacaaccacc 19
||||||| | |||| |||
Db 26 GTGCACGCAAACACGGAACGC 8

RESULT 11
AZ492829/c
LOCUS AZ492829 34 bp DNA GSS 05-OCT-2000
DEFINITION IM0427.f 2 Mouse 10kb plasmid UniZAP library Mus musculus, genomic clone UNIZAPIM0427.f11 2', DNA sequence.
ACCESSION AZ492829
VERSION AZ492829.1 GI:10665614
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 34)
AUTHORS Dunn,D., Aoyama,A., Barber,M., Beierom,L., Burrill,L., Barrett, Islam,R., Lomascro,S., Mahmood,M., Meemen,E., Peterson,L., Petty,

pIN13 vector. Library is normalized. Library was
constructed by Beatriz Soares and M. Fatima Bonaldo. "
BASE COUNT 6 a 14 c 20 g 12 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 10; Length 52;
Best Local Similarity 82.4%; Pred. No. 1.2e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 cgccccacagcccg 20
|||||||||| ||| |
Db 10 CAGCCCACAGCCAG 33

RESULT 14
AI439245
LOCUS AI439245 52 bp mRNA EST 03-MAR-1999
DEFINITION ty52a06.x1 NCI_CGAP_Lym12 Homo sapiens cDNA clone IMAGE:2134738 3'
similar to SW:PRP4_HUMAN P04280 SALIVARY PROLINE RICH PROTEIN PO
PRECURSOR contains MER22.1 MSR1 repetitive element ;, mRNA
sequence.
ACCESSION AI439245
VERSION AI439245.1 GI:3430272
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 52)
AUTHORS NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Life Technologies cation #: 11547-015
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI_CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 2765 Std Error: 0.00
Seq primer: 4000 from 1ibco
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1..52
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2134738"
/clone_lib="NCI_CGAP_Lym12"
/tissue_type="lymphoma, follicular mixed small and large
cell"
/lab_host="DH10B"
/note="organ: lymph node; Vector: pCMV SPORT6; Site 1:
Sall; Site 2: NotI; Cloned unidirectionally. Primer:
oligo dT. Average insert size 1.25 kb. Life Technologies
catalog #: 11547-015"
BASE COUNT 18 a 21 c 4 g 9 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 10; Length 52;
Best Local Similarity 82.4%; Pred. No. 1.2e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ccgccccacagcccgc 19
||| ||| ||||| |||
Db 9 CCAACCCCCAGGAACGC 19

RESULT 15
BG036071/c
LOCUS BG036071 54 bp mRNA EST 21-JAN-2...
DEFINITION 602326294T1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4414.46 3'
mRNA sequence.
ACCESSION BG036071
VERSION BG036071.1 GI:12430839
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE 1 (bases 1 to 54)
AUTHORS NIH_MGC http://www.ncbi.nlm.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10141 row: I column: 23
High quality sequence stop: 54.
FEATURES Location/Qualifiers
source 1..54
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4414246"
/clone_lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage resistant)"
/note="organ: liver; Vector: pCMV SPORT6; Site 1:
Site 2: Sall; Cloned unidirectionally; cloned at primed;
Average insert size 1.7 kb. Library enriched for
full length clones and constructed by Life Technologies
Note: this is a NIH MGC Library."
BASE COUNT 11 a 9 c 20 g 14 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 11; Length 54;
Best Local Similarity 82.4%; Pred. No. 1.2e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 cgccccacagcccgc 20
|||||||||| ||| | |
Db 52 CAGCCCACAAGACCGAC 36

Search completed: February 9, 2002, 01:01:55
Job time: 21881 sec

Compare version 4.5

Run mode: nucleic search, using sw model

Run on: February 9, 2002, 01:44:07 ; Search time 25.19.39 Seconds
(without alignments)
130.962 Million cell updates/sec

Title: US-10-006-430-71
Perfect score: 20
Sequence: 1 ggccgcaaatgcccctcg 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 824806975 residues

Total number of hits satisfying chosen parameters: 586446

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_or:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

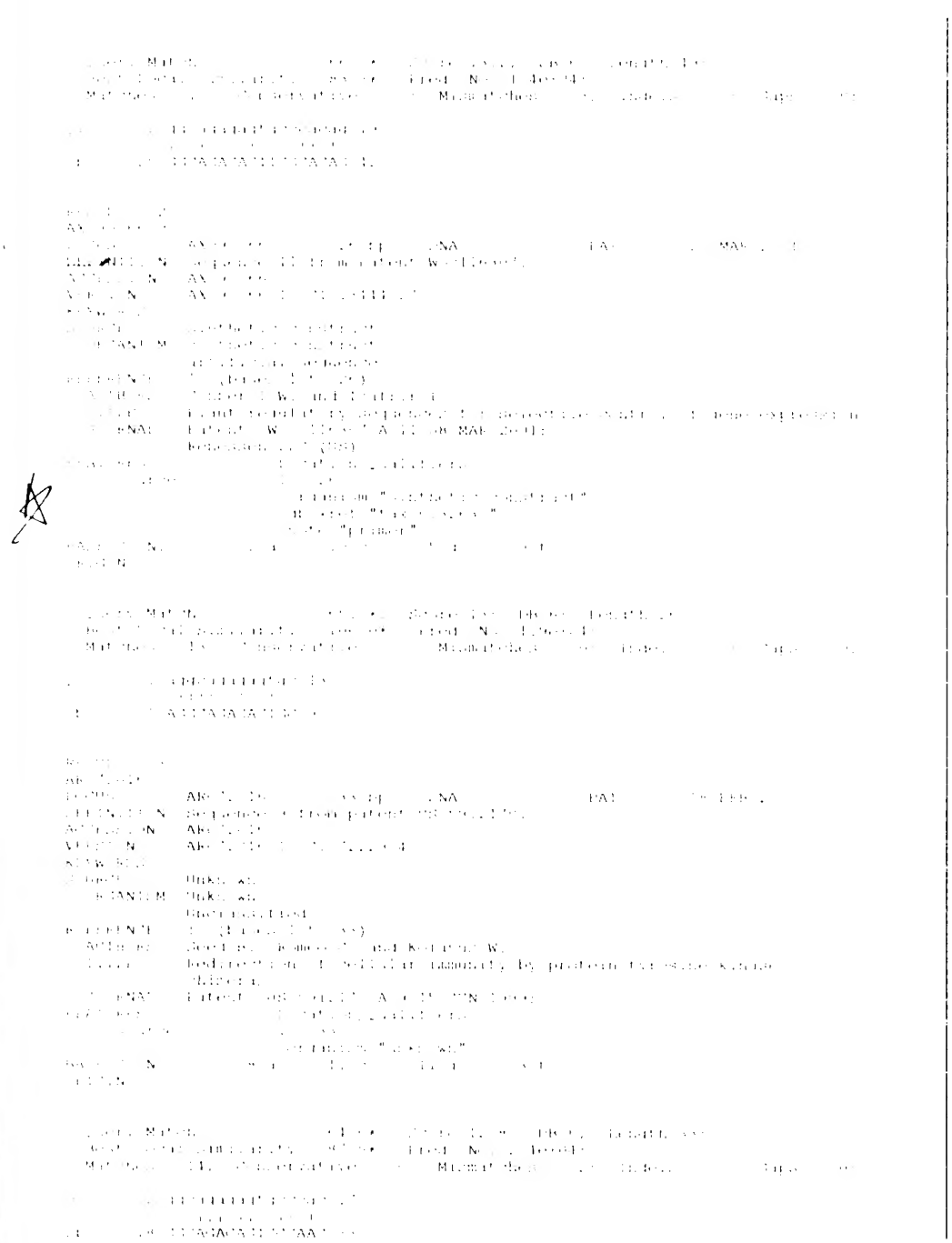
Prob. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	13.2	66.0	49	6	AR104156	AR104156 Sequence
c 2	13	65.0	26	6	AX090066	AX090066 Sequence
3	12.8	64.0	33	6	AR072016	AR072016 Sequence
4	12.8	64.0	34	6	AR072014	AR072014 Sequence
5	12.8	64.0	34	6	AR095772	AR095772 Sequence
c 6	12.8	64.0	51	6	AX157425	AX157425 Sequence
7	12.6	63.0	43	6	AX166824	AX166824 Sequence
8	12.4	62.0	51	6	AX160036	AX160036 Sequence
9	12.2	61.0	24	6	AR104154	AR104154 Sequence
c 10	12.2	61.0	28	6	170267	170267 Sequence 6
11	12.2	61.0	31	6	AR065647	AR065647 Sequence
c 12	12.2	61.0	34	6	AR151524	AR151524 Sequence
13	12.2	61.0	39	6	AX011250	AX011250 Sequence
14	12.2	61.0	51	6	AX162227	AX162227 Sequence
15	12.2	61.0	51	6	AX165008	AX165008 Sequence
16	12	60.0	30	6	AR112370	AR112370 Sequence
17	12	60.0	43	6	AR050692	AR050692 Sequence
18	12	60.0	46	6	134406	134406 Sequence 5
c 19	12	60.0	54	6	A81699	A81699 Sequence 4
20	12	60.0	56	6	AX146594	AX146594 Sequence
c 21	12	60.0	58	6	AX146595	AX146595 Sequence
22	12	60.0	59	6	AX011333	AX011333 Sequence
c 23	12	60.0	60	5	CHKREV11	M2134 Saitis stati
c 24	11.8	59.0	19	6	AX035693	AX035693 Sequence
c 25	11.8	59.0	26	6	AX005966	AX005966 Sequence
26	11.8	59.0	36	6	A61609	A61609 Sequence 4
c 27	11.8	59.0	40	9	BSEREG3	BSEREG3 Bombyx sapien
c 28	11.8	59.0	42	6	A87736	A87736 Sequence 30
c 29	11.8	59.0	42	6	A87766	A87766 Sequence 60
30	11.8	59.0	48	6	A61610	A61610 Sequence 5
31	11.8	59.0	51	6	A61633	A61633 Sequence 28
c 32	11.8	59.0	51	6	AX162443	AX162443 Sequence
c 33	11.8	59.0	51	6	AX162444	AX162444 Sequence
34	11.8	59.0	54	6	A61632	A61632 Sequence 29
c 35	11.6	58.0	20	6	AR122518	AR122518 Sequence
36	11.6	58.0	20	6	AX000378	AX000378 Sequence
c 37	11.6	58.0	20	6	E14563	E14563 Pk protein
38	11.6	58.0	21	6	AX095004	AX095004 Sequence
c 39	11.6	58.0	26	6	106781	106781 Sequence 7
40	11.6	58.0	27	6	AR143903	AR143903 Sequence
c 41	11.6	58.0	27	6	107589	107589 Sequence 7
c 42	11.6	58.0	31	6	E32339	E32339 Spe-1-08 spe
c 43	11.6	58.0	32	6	AR140100	AR140100 Sequence
c 44	11.6	58.0	32	6	E15872	E15872 Pk protein
45	11.6	58.0	34	6	AR085000	AR085000 Sequence

ALIGNMENTS

RESULT 1
AR104156
LOCUS AR104156 49 bp DNA 01-FEB-2000
DEFINITION Sequence 5 from patent US 6093542.
ACCESSION AR104156
VERSION AR104156.1 GI:12816864
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 49)
AUTHORS Romano,J., Pal,R. and Shuttleff,R.
TITLE Isothermal transcription based amplification assay for the detection and quantification of macrophage derived chemokine mRNA
JOURNAL Patent: US 6093542 A 5 25-JUL-2000;
FEATURES Location/Qualifiers
source 1..49
/organism="unknown"
BASE COUNT 14 a 10 c 15 g 10 t
ORIGIN



```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521 A 754 07 JUN 2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (754 is other entry)
                     Accession number cq2425687"
BASE COUNT         9 a      18 c      16 g       8 t
ORIGIN


  Query Match             64.0%;  Score 12.8;  DB 6;  Length 51;
  Best Local Similarity   87.5%;  Prob. No. 2.4e-04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 gcagagactgccca 16
           ||||||||||||||
Db      27 GCAGAGACTGCCAC 40


RESULT   7
AX166824
LOCUS       AX166824          43 bp    DNA             PAT       03-JUL-2001
DEFINITION  Sequence 1 from Patent WO0144567.
ACCESSION   AX166824
VERSION     AX166824.1  GI:14506427
KEYWORDS    .
SOURCE      Human immunodeficiency virus.
  ORGANISM  Human immunodeficiency virus
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1  (bases 1 to 43)
  AUTHORS   Bonetjoe,A.R., Becker,M.M., Browne,E.A., Friedenberg,M.C.,
            Bajjar,F.F., Laytom,A., Menou,L. and Tord,C.
  TITLE     Process for labelling a nucleic acid
  JOURNAL   Patent: WO 0144507 A 1 21 JUN 2001;
            Bio MERIEUX (FR) ; Gen Probe Incorporated (US)
FEATURES             Location/Qualifiers
     source          1..43
                     /organism="Human immunodeficiency virus"
                     /db_xref="taxon:12721"
BASE COUNT        12 a      15 c       7 g       4 t
ORIGIN


  Query Match             64.0%;  Score 12.8;  DB 6;  Length 43;
  Best Local Similarity   78.6%;  Prob. No. 3.1e-04;
  Matches   11;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 tcgcagactgccaca 15
           ||| | | || |||||
Db      12 ACGGAGACAGAGCAACAC 27


RESULT   8
AX166049
LOCUS       AX166049          51 bp    DNA             PAT       22-JUN-2001
DEFINITION  Sequence 4364 from Patent WO0140521.
ACCESSION   AX166049
VERSION     AX166049.1  GI:14541967
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521 A 4364 07 JUN 2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (4364 is other entry)
                     Accession number cq4425687"
BASE COUNT        14 a      14 c      14 g       9 t
ORIGIN


  Query Match             62.0%;  Score 12.4;  DB 6;  Length 51;
  Best Local Similarity   92.9%;  Prob. No. 4e-04;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 gcagagactgccca 16
           ||||||||||||| ||
Db      27 GCAGAGACTGCCAC 40


RESULT   9
AR104154
LOCUS       AR104154          24 bp    DNA             PAT       14-FEB-2001
DEFINITION  Sequence 4 from patent US 6093542.
ACCESSION   AR104154
VERSION     AR104154.1  GI:12816862
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Romano,J., Pal,K. and Shartlett,K.
  TITLE     Isothermal transcription based amplification assay for the
            detection and quantitation of macrophage derived chemokine RNA
  JOURNAL   Patent: US 6093542-A 4 25 JUL 2000;
FEATURES             Location/Qualifiers
     source          1..24
                     /organism="unknown"
BASE COUNT         5 a       5 c       9 g       5 t
ORIGIN


  Query Match             61.0%;  Score 12.2;  DB 6;  Length 24;
  Best Local Similarity   82.4%;  Prob. No. 5.6e-04;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 gcagagactgccaca 16
           ||||||:|  ||||
Db       1 GCAGAGACTTGGCACAG 17


RESULT   10
170267/c
LOCUS       170267            28 bp    DNA             PAT       15-AUG-2000
DEFINITION  Sequence 5 from patent US 5679525.
ACCESSION   170267
VERSION     170267.1  GI:3006402
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 28)
  AUTHORS   Peterson,M.Gregory and Beskett.
  TITLE     Epstein barr virus transcription factor binding us. 6
  JOURNAL   Patent: US 5679525 A 5 21 OCT 1997;
FEATURES             Location/Qualifiers
     source          1..28
```

```
LOCUS       AX105008        51 bp    DNA             PAT       22-JUN-2001
DEFINITION  Sequence 203 from Patent WO 0138586.
ACCESSION   AX105008
VERSION     AX105008.1  GI:14545847
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0138586 A 203 31-MAY-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers.
     source          1..51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     variation       25
                     /note="single nucleotide polymorphism
                     Accession number cp44524222"
BASE COUNT       13 a       7 c      24 g       7 t
ORIGIN


  Query Match             51.0%;  Score 12.2;  DB 6;  Length 51;
  Best Local Similarity   82.4%;  Pred. No. 5.2e-04;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 AAAAAAAAAAAAAAC 17
           ||| |||||||| ||||
Db      20 AAGAUGAGAGUGACCAC 45



Search completed: February  6, 2002, 01:44:08
Job time: 1897 sec
```

GenScore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

SW nucleic - nucleic search, using sw model

Run on: February 9, 2002, 01:02:16 ; Search time 4962.23 Seconds
(without alignments)
43.410 Million cell updates/sec

Title: US-10-006-430-75

Perfect score: 20

Sequence: 1 tctgattacagttaaaaaaa 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1145947 seqs, 567280741 residues

Total number of hits satisfying chosen parameters: 111874

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estrom:*
2: em_esthum:*
3: em_estmus:*
4: em_estoms:*
5: em_estpln:*
6: em_estbac:*
7: em_estvrt:*
8: em_estunc:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_hum:*
15: em_gss_hum:*
16: em_gss_mus:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_vrt:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	14.8	69.0	47	13	AZ940264	AZ940264 2M0199003
2	14.8	69.0	58	13	AZ652535	AZ652535 1M0525K18
3	13.6	68.0	34	13	AZ413987	AZ413987 1M0188D15
4	13	65.0	39	13	AZ393934	AZ393934 1M0328F19
5	12.6	63.0	34	10	AA906810	AA906810 ok73e02.s
6	12.6	63.0	54	13	AZ413567	AZ413567 1M0197004
7	12.4	62.0	57	11	C01641	C01641 BUMGS000865
8	12.4	62.0	57	11	H47597	H47597 yp75d02.s1
9	12.4	62.0	59	13	AZ650142	AZ650142 1M0520110
10	12.2	61.0	46	11	H08749	H08749 y18d03.s1
11	12.2	61.0	58	10	AW460588	AW460588 da26d05.x
12	12	60.0	56	11	H77417	H77417 BS077417 Ba
13	12	60.0	58	10	AA422138	AA422138 zm45b12.r1
14	12	60.0	59	10	AA106633	AA106633 zm23d11.r1
15	12	60.0	60	10	AA704865	AA704865 zm55f12.s1
16	11.8	59.0	27	13	AZ764970	AZ764970 1M0741112
17	11.8	59.0	34	10	A1194779	A1194779 qz50d12.x
18	11.8	59.0	37	11	BE122604	BE122604 601700622
19	11.8	59.0	53	11	H18867	H18867 zm45b12.r1
20	11.6	58.0	28	11	H63106	H63106 yt48a04.s1
21	11.6	58.0	31	10	AV832900	AV832900 AV832900
22	11.6	58.0	32	11	H70643	H70643 yu18b03.s1
23	11.6	58.0	40	10	AA581049	AA581049 to92d11.r1
24	11.6	58.0	50	10	AU102648	AU102648 AU102648
25	11.6	58.0	50	10	AU105833	AU105833 AU105833
26	11.6	58.0	50	10	AU105874	AU105874 AU105874
27	11.6	58.0	50	10	AU105875	AU105875 AU105875
28	11.6	58.0	50	10	AU105876	AU105876 AU105876
29	11.6	58.0	50	10	AU105877	AU105877 AU105877
30	11.6	58.0	50	10	AU105878	AU105878 AU105878
31	11.6	58.0	50	10	AU105879	AU105879 AU105879
32	11.6	58.0	50	10	AU105882	AU105882 AU105882
33	11.6	58.0	50	10	AU105884	AU105884 AU105884
34	11.6	58.0	50	10	AU105886	AU105886 AU105886
35	11.6	58.0	55	10	AA469546	AA469546 zt27d07.r1
36	11.6	58.0	60	11	BE638499	BE638499 NL3sp04
37	11.4	57.0	31	13	AZ798032	AZ798032 2M0154C
38	11.4	57.0	32	13	TA145830	A147071 1 Bacce
39	11.4	57.0	43	13	AZ327966	AZ327966 1M0511E12
40	11.4	57.0	51	10	A1280432	A1280432 qt45c11.x
41	11.4	57.0	52	13	B01670	B01670 09K 13d7
42	11.4	57.0	56	11	C00688	C00688 BUMGS000824
43	11.2	56.0	25	13	AZ371108	AZ371108 1M0122D03
44	11.2	56.0	28	13	AZ438215	AZ438215 1M0228K12
45	11.2	56.0	36	13	AZ807406	AZ807406 2M0070D15

ALIGNMENTS

RESULT 1
AZ940264
LOCUS AZ940264 47 bp DNA GSS 2? APR 2002
DEFINITION 2M0199003R Mouse 10kb plasmid UBC?2M library Mus musculus genomic clone UBC?2M0199003 R, DNA sequence.
ACCESSION AZ940264
VERSION AZ940264.1 GI:14801605
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
REFERENCE 1 (bases 1 to 47)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beaudry,L., Duval,B., Ramos,C., Islam,R., Longacre,S., Mahmood,M., Meehan,E., Pedersen,L., Berti,M., Roso,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dbunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0199 row: 0 column: 03
Seq primer: CACACAGGAAACAGCTATGAC
Class: plasmid ends
High quality sequence stop: 47.
FEATURES location/qualifiers
source 1..47

```
FEATURES             Location/Qualifiers
     source          1..34
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UBGC1M003028E15"
                     /clone_lib="Mouse 10kb plasmid UBGC1M library"
                     /sex="Male"
                     /lab_host="E. coli strain XL10-Gold, Tet resistant, F'"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/).  The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity.  The sheared DNA
                     was blunt-end repaired with T4 DNA polymerase and T4
                     polynucleotide kinase.  Adaptor oligonucleotides were
                     ligated to the blunt ends in high molar excess.  The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis.  Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy number
                     inducible derivative of plasmid R1.  The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified.  The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        14 a        8 c        6 g       14 t
ORIGIN


  Query Match             68.0%;  Score 13.6;  DB 14;  Length 34;
  Best Local Similarity   80.0%;  Pred. No. 1.2e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 tgtgattacaattcaaaaaaaaaaaaaaaaaa 20
           ||| |||| ||||| ||| |
Db      11 TGTGATTAAACTTGTAGCGTG 30


RESULT   4
AA906844
ID   AA906844       standard; DNA; EST; 39 BP.
XX
AC   AA906844;
XX
DT   05-OCT-2000 (Rel. 67, Created)
DT   05-OCT-2000 (Rel. 67, Last updated, Version 1)
XX
DE   UBGC1M003028E19.g1 Mouse 10kb plasmid UBGC1M library Mus musculus genomic
DE   clone UBGC1M003028E19 g1, DNA sequence.
XX
KW   EST.
XX
OS   Mus musculus
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
XX
RN   [1]
RP   1-39
RA   Dong Y., Acevedo,A., Barker,M., Beerman,I., Bunaf,B., Bunil,C.,
RA   Eslam,H., Lambacher,S., Mahmoud,M., Meenen,E., Pedersen,L., Reilly
RA   ,M., Rose,M., Rose,R., Stokes,K., Tingey,A., von Niederhausern,A.
RA   and Wright,G.;Weiss,R.
RT   "Mouse whole genome scaffolding with paired end reads from 10kb
RT   plasmid inserts"
RL   Unpublished (2000)
XX
RN   [2]
RA   Contact: Robert W. Weiss
RA   University of Utah Genome Center
RA   University of Utah
RA   bmw, 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
RA   84112, USA
RA   Tel: 801 585 3566
RA   Fax: 801 585 7122
RA   Email: dbmail@genetics.utah.edu
RA   Insert Length: 10000  Std Error: 0.00
RA   Plate: 0328  row: F  column: 19
RA   Seq primer: CACACAGGAAACAGCTATGAC
RA   Class: plasmid ends
```

```
                     High quality sequence stop: 39.
FEATURES             Location/Qualifiers
     source          1..39
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UBGC1M003028E19"
                     /clone_lib="Mouse 10kb plasmid UBGC1M library"
                     /sex="Male"
                     /lab_host="E. coli strain XL10-Gold, Tet resistant, F'"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/).  The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity.  The sheared DNA
                     was blunt-end repaired with T4 DNA polymerase and T4
                     polynucleotide kinase.  Adaptor oligonucleotides were
                     ligated to the blunt ends in high molar excess.  The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis.  Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy number
                     inducible derivative of plasmid R1.  The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified.  The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        10 a        8 c       14 g       14 t
ORIGIN


  Query Match             65.0%;  Score 13;  DB 14;  Length 39;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 tgtgattacaatt 13
           |||||||||||||
Db      23 TGTGATTACAATT 35


RESULT   5
AA906810
ID   AA906810       standard; DNA; EST; 34 BP.
XX
AC   AA906810;
XX
DT   05-OCT-2000 (Rel. 67, Created)
DT   05-OCT-2000 (Rel. 67, Last updated, Version 1)
XX
DE   ok73e02.s1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1675877 3'
DE   similar to SW:UFD1_HUMAN Q92890 UBIQUITIN FUSION DEGRADATION
DE   PROTEIN 1 HOMOLOG ;, mRNA sequence.
XX
KW   EST.
XX
OS   Homo sapiens
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
XX
RN   [1]
RP   1-34
RA   NCI, CGAP Project;
RT   "National Cancer Institute, Cancer Genome Anatomy Project (CGAP)
RT   Tumor Gene Index"
RL   Unpublished (1997)
XX
CC   Contact: Robert Strausberg, Ph.D.
CC   Email: cgapbs-r@mail.nih.gov
CC   Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D.; Mitchell
CC   Emmert-Buck, M.D., Ph.D.
CC   cDNA Library Preparation: M. Benito Soares, Ph.D.
CC   cDNA Library Arrayed by: Greg Lennon, Ph.D.
CC   DNA Sequencing by: Washington University Genome Sequencing Center
CC   Clone distribution: NCI CGAP clone distribution info and clone can be
CC   found through the I.M.A.G.E. Consortium/LLNL at
CC   www.bio.llnl.gov/bbrp/image/image.html
XX
CC   Trace considered overall poor quality
CC   Seq primer:  40ml4 fwd, EL from Amersham
```

```
..        2 attattacaat ga 15
            ||||||||||| |
5'      11 GTGCATTACAGTAAA 42


RESULT   8
AI457...
ID   AI457...       447 bp    mRNA            EST       16 AUG 1999
DEFINITION  yp75d02.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:194251 3', similar to gb:L07077 ENOYL-CoA HYDRATASE (HUMAN);,
            mRNA sequence.
ACCESSION   AI457...
VERSION     AI457...1  GI:2...3640
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Bailler T., Clark,N., Dutcupe,T., Ellston,K., Hawkins,M., Holman
            ,M., Hoffman,M., Kucaba,T., Lee,M., Lennon,G., Marra,M., Parsons,J.,
            Pitkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The Washu-M...-K EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 3...286 1800
            Fax: 3...286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1450
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered ok and not poor quality
            Insert length: 1450   Std Error: 0.00
            Seq primer: Promega  z1ml4
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1..447
                     /organism="Homo sapiens"
                     /db_xref="GDB:...6222"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:194251"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I   oligo(dT) primer
                     [5' AACTGGAACAATTAATTAAAGATC(T)17 3'],
                     double stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Benito Soares and M.Fatima Bonaldo."
BASE COUNT        11 a      15 c      15 g      14 t       2 others
ORIGIN


  Query Match             62.0%;  Score 12.4;  DB 11;  Length 57;
  Best Local Similarity   92.9%;  Prod. No. 4.7e-04;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 tgtgattacaattg 14
           || | ||||||||||
Db      41 TGCGATTACAGTTG 54


RESULT   9
AZ650142/c
LOCUS       AZ650142         59 bp    DNA             GSS       11-...-...
DEFINITION  IM0520110F Mouse 10kb plasmid 000C1M library Mus musculus ge...mic
            clone 000C1M0520110 F, DNA sequence.
ACCESSION   AZ650142
VERSION     AZ650142.1  GI:11784329
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beaverm,L., Duval,B., Hurd,T.,
            Islam,R., Londarte,S., Mahmood,M., Mcneil,E., Pedersen,L., Reilly
            ,M., Rose,M., Rose,R., Stokes,K., Tingey,A., von Niederhausern A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10 kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert M. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 408, Biomedical Polymers Research Bldg., 20 S. 2...0 E., SLC,
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: rdunn@genetics.utah.edu
            Insert length: 10000   Std Error: 0.00
            Plate: 0520  row: I  column: 10
            Seq primer: CGTTGIAAAACGACGGCCAGI
            Class: plasmid ends
            High quality sequence stop: 59.
FEATURES             Location/Qualifiers
     source          1..59
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="000C1M0520110"
                     /clone_lib="Mouse 10kb plasmid 000C1M library."
                     /sex="Male"
                     /lab_host="E. Coli strain XL10 Gold, ultracompetent, [ "
                     /note="Vector: PWD42nn; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/cmr/pro5/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high molar excess. The
                     adaptored DNA was purified and size-selected [ ] a 9.5
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4734214|gb|AF129072.1), a copy number
                     inducible derivative of plasmid RI. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed [ ]
                     chemically-competent E. coli XL10 Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        18 a      12 c      11 g      18 t
ORIGIN


  Query Match             62.0%;  Score 12.4;  DB 15;  Length 59;
  Best Local Similarity   92.9%;  Prod. No. 4.7e-04;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 tgtgattacaattg 14
           || | ||||||||||
Db      15 TGTGCTTACAGTTG 2...
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 56)
  AUTHORS   Vogt,T. and Mecklenburg,M.
  TITLE     UV B and IFA regulated gene expression in human melanocyte
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Vogt, Thomas
            Med. Biol.
            Sidney Kimmel Cancer Center
            3960 Sorrento Park Road, Ste. 200, San Diego, CA 92121, USA
            Email: DRVOGT@AOL.COM
            UV is inducible.
FEATURES             Location/Qualifiers
     source          1..56
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DVIN7"
                     /clone_lib="Human newborn melanocytes (T.Vogt)"
                     /cell_type="melanocytes"
                     /cell_line="Clonetics Corp. (San Diego, CA) strain #68 and
                     2486"
                     /dev_stage="newborn"
BASE COUNT        14 a        8 c        16 g        18 t
ORIGIN


  Query Match             60.0%;  Score 12;  DB 11;  Length 56;
  Best Local Similarity   75.0%;  Pred. No. 7.4e+04;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 tgtgattacagttaaaagcg 20
           ||||| ||||| | || |
Db      42 TGTGATACAGTGAATGGTG 61


RESULT  13
AA122108/c
LOCUS       AA122108          56 bp    mRNA             EST       16-OCT-1997
DEFINITION  zz43d03.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
            IMAGE:755247 5', mRNA sequence.
ACCESSION   AA122108
VERSION     AA122108.1  GI:2105066
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 56)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
            , Schellenberg,K., Steptoe,M., Tan,F., Theisinq,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
  TITLE     WashU Merck EST Project 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Initiative full length read
            The vector to vector length is 190
            Seq primer:  zRmB54 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1..58
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:755247"
                     /clone_lib="Soares ovary tumor NbHOT"
                     /sex="Female"
                     /tissue_type="ovarian tumor"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: ovary; Vector: pT7T3D (Pharmacia) w/ 5.a
                     modified polylinker; Site 1: Not I; Site 2: Eco RI 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer (5'
                     PGEIA'CAATTCTCAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTT 3'),
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pUC vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo. "
BASE COUNT        21 a        14 c        9 g        14 t
ORIGIN


  Query Match             60.0%;  Score 12;  DB 11;  Length 58;
  Best Local Similarity   100.0%;  Pred. No. 7.4e+04;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 tgtgattacagt 12
           ||||||||||||
Db      12 TGTGATACAGTC 1


RESULT  14
AA106633/c
LOCUS       AA106633          59 bp    mRNA             EST       04-FEB-1997
DEFINITION  nma18b05.r1 Stratagene mouse diaphragm (#937304) Mus musculus cDNA
            clone IMAGE:521841 5', mRNA sequence.
ACCESSION   AA106633
VERSION     AA106633.1  GI:1656415
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,G., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Waterston,R.,
            Theisinq,B., Wylie,T., Lennon,G., Soares,B., Wilcox,K. and
            Waterston,R.
  TITLE     The WashU HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU HHMI Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouse@watson.wustl.edu
            This clone is available royalty free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:415689
            Seq primer:  zRmB54 rev1 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1..59
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:521841"
                     /clone_lib="Stratagene mouse diaphragm (#937304)"
                     /tissue_type="diaphragm"
                     /dev_stage="adult"
                     /lab_host="SOLK (kanamycin resistant)"
                     /note="Organ: diaphragm; Vector: pBluescript SK -; Site 1:
                     EcoRI; Site 2: XhoI; Cloned unidirectionally with mRNA
                     prepared from diaphragm muscle.  Primer: oligo dT Acetoam
                     insert size: 1.5 kb, Gnl ZAP XR Vector  5' adaptor
                     sequence: 5' GAATTCGGCACGAG 3'  3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTT 3'"
BASE COUNT        12 a        12 c        19 g        16 t
ORIGIN
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: February 9, 2002, 06:24:00 ; Search time 1864.18 Seconds
(without alignments)
177.086 Million cell updates/sec

Title: US-10-006-430-76
Perfect score: 20
Sequence: 1 acgaatcaacgatgttgtgca 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 147240 seqs, 82485079 residues

Total number of hits satisfying chosen parameters: 586436

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_or:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	13.8	69.0	48	6	AX076651	AX076651 Sequence
2	13.8	69.0	51	6	AX164881	AX164881 Sequence
3	13.2	66.0	28	6	E05487	E05487 PCR primer
4	13.2	66.0	51	6	AX164880	AX164880 Sequence
c 5	13	65.0	42	6	107146	107146 Sequence 7
6	12.8	64.0	24	6	E16132	E16132 Sequence 24
c 7	12.8	64.0	27	6	157990	157990 Sequence 5
8	12.8	64.0	29	6	E26920	E26920 Mutant seq
c 9	12.8	64.0	30	6	AR064721	AR064721 Sequence
c 10	12.8	64.0	30	6	AR089162	AR089162 Sequence
11	12.8	64.0	30	6	AR153238	AR153238 Sequence
c 12	12.8	64.0	60	6	105845	105845 Sequence 8
13	12.6	63.0	57	6	AX141071	AX141071 Sequence
14	12.6	63.0	57	6	AX146941	AX146941 Sequence
15	12.6	63.0	60	10	MUS1GHCAA	MF0879 Mouse 14 se
16	12.6	63.0	60	10	MUS1GHTH1	MF0880 Mouse 14
c 17	12.4	62.0	21	6	AR049055	AR049055 Sequence
18	12.4	62.0	21	6	AX095333	AX095333 Sequence
19	12.4	62.0	58	6	A25701	A25701 DC PVII INK
20	12.4	62.0	58	6	AR051374	AR051374 Sequence
21	12.4	62.0	60	6	E22145	E22145 Yeast expat
22	12.4	62.0	41	6	AR090075	AR090075 Sequence
c 23	12.2	61.0	34	6	AR142300	AR142300 Sequence
c 24	12.2	61.0	34	6	127173	127173 Sequence 28
c 25	12.2	61.0	34	6	132754	132754 Sequence 28
c 26	12.2	61.0	34	6	184688	184688 Sequence 28
c 27	12.2	61.0	35	6	AX078922	AX078922 Sequence
28	12.2	61.0	36	6	AR142301	AR142301 Sequence
29	12.2	61.0	36	6	127174	127174 Sequence 29
30	12.2	61.0	36	6	132755	132755 Sequence 29
31	12.2	61.0	36	6	184689	184689 Sequence 29
c 32	12.2	61.0	51	6	AX162560	AX162560 Sequence
33	12.2	61.0	51	6	AX164882	AX164882 Sequence
c 34	12	60.0	22	6	AR008881	AR008881 Sequence
c 35	12	60.0	22	6	AR087576	AR087576 Sequence
c 36	12	60.0	22	6	185132	185132 Sequence 6
37	12	60.0	25	6	AR124224	AR124224 Sequence
38	12	60.0	26	6	AX148206	AX148206 Sequence
c 39	12	60.0	27	6	AX183818	AX183818 Sequence
c 40	12	60.0	28	9	HUMD2D03M5	D16905 Human Hep32
c 41	12	60.0	28	9	HUMD6H06M5	D17115 Human Hep32
42	12	60.0	30	6	A48425	A48425 Sequence 48
43	12	60.0	30	6	AX083705	AX083705 Sequence
44	12	60.0	49	6	AX183593	AX183593 Sequence
45	12	60.0	51	6	AX190347	AX190347 Sequence

ALIGNMENTS

RESULT 1
AX076651
LOCUS AX076651 48 bp DNA PAT 09-FEB-2002
DEFINITION Sequence 167 from Patent WO0103719.
ACCESSION AX076651
VERSION AX076651.1 GI:12711188
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 48)
AUTHORS Boyle,W.J., Lacey,D.L., Calzone,F.J., Chang,M.S. and Senaldi,G.
TITLE Combination therapy for conditions leading to bone loss
JOURNAL Patent: WO 0103719-A 167 18-JAN-2001;
Amgen Inc. (US)
FEATURES Location/qualifiers
source 1..48
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="PCR primer for deletion mutant "

```
    JOURNAL   Patent: EP 0541802-A1 5-15-NOV-1992;
FEATURES             Location/Qualifiers
     source          1..42
                     /organism="unknown"
BASE COUNT        12 a      10 c       5 g       5 t
ORIGIN


  Query Match             65.0%;  Score 13;  DB 6;  Length 42;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 aatcaggatatgg 17
           |||||||||||||
Db      18 AGTCAGGATGTTC 6


RESULT   6
E16132
LOCUS       E16132          24 bp    DNA             PAT        28-JUL-1999
DEFINITION  Sequence of Eos 1 binding motif.
ACCESSION   E16132
VERSION     E16132.1  GI:971005
KEYWORDS    JP 1998137000 A/1.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Taniguchi,N.
  TITLE     SCREEN NO OF METASTASIS SUPPRESSOR
  JOURNAL   Patent: JP 1998137000 A 1 26-MAY-1998;
            SUNTORY LTD, TANIGUCHI NAOYUKI
COMMENT     OS   None
            OC   Artificial sequences.
            PN   JP 1998137000 A/1
            PD   26-MAY-1998
            PF   01-NOV-1996 JP 1996305486
            PI   TANIGUCHI NAOYUKI
            PC   C12Q1/68;C07H21/04;C07K13/00;G01N33/50;G01N33/53;C07K14  ;C12N15/09;
            CC   Sequence Structure: Single;
            CC   Topology: Linear;
            CC   Hypothetical: No;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1..24
            FT                   /organism="Artificial sequences".
FEATURES             Location/Qualifiers
     source          1..24
                     /organism="unidentified"
                     /db_xref="taxon:42644"
BASE COUNT         7 a       0 c       13 g       4 t
ORIGIN


  Query Match             64.0%;  Score 12.8;  DB 6;  Length 24;
  Best Local Similarity   87.5%;  Pred. No. 1.8e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 aggatcaggatgtgt 18
           ||||| |||||| |||
Db       2 AGGTCAGGATGAGTG 17


RESULT   7
AF2990/c
LOCUS       AF2990          27 bp    DNA             PAT        07-OCT-1997
DEFINITION  Sequence 9 from patent US 5610147.
ACCESSION   AF2990
VERSION     AF2990.1  GI:2483054
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.

REFERENCE   1  (bases 1 to 27)
  AUTHORS   Townes,T.M. and McCune,S.L.
  TITLE     Transgenic, cross-linked hemoglobin
  JOURNAL   Patent: US 5610147-A 6 11-MAR-1997;
FEATURES             Location/Qualifiers
     source          1..27
                     /organism="unknown"
BASE COUNT         9 a       9 c       5 g       4 t
ORIGIN


  Query Match             64.0%;  Score 12.8;  DB 6;  Length 27;
  Best Local Similarity   87.5%;  Pred. No. 1.8e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 aggatcaggatgtgt 18
           ||||||| ||| |||
Db      23 GGAGTCAGGATGAGTGT 8


RESULT   8
E26920
LOCUS       E26920          29 bp    DNA             PAT        07-DEC-1999
DEFINITION  Mutant secretory device enzyme capable of efficiently secreting
            protein into medium in coryneform bacterium.
ACCESSION   E26920
VERSION     E26920.1  GI:4026840
KEYWORDS    JP 1999169182 A/11.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 29)
  AUTHORS   Takashi,K.,Y.A.A. and Yukawa.
  TITLE     Mutant secretory device enzyme capable of efficiently secreting
            protein into medium in coryneform bacterium
  JOURNAL   Patent: JP 1999169182 A 11 29-JUN-1999;
            MITSUBISHI CHEM CORP
COMMENT     OS   Unidentified
            PN   JP 1999169182 A/11
            PD   29-JUN-1999
            PF   10-DEC-1997 JP 1997361768
            PR
            PI   TAKASHI KOBAYASHI,YUKO ASAI,HIDEAKI YUKAWA
            PC   C12N15/09,C12N1/21,C12P  ,  ;C12N15/09,C12R1:19),(C12N1/21,
            PC   C12R1:19),
            PC   (C12N1/21,C12R1:19),(C12P21/02,C12R1:19),C12N15/  ,C12N15/09,
            PC   C12R1:19)
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1..29
            FT                   /organism="Unidentified"
FEATURES             Location/Qualifiers
     source          1..29
                     /organism="unidentified"
                     /db_xref="taxon:42644"
BASE COUNT         6 a       7 c       7 g       9 t
ORIGIN


  Query Match             64.0%;  Score 12.8;  DB 6;  Length 29;
  Best Local Similarity   87.5%;  Pred. No. 1.8e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 acggatcaggatgtg 16
           |||||||||| || |||
Db      14 ACGGAGTCAGAATGTT 29


RESULT   9
AR064721/c
LOCUS       AR064721        30 bp    DNA             PAT        27-SEP-1999
DEFINITION  Sequence 37 from patent US 5864306.
```

```
  Query Match             63.0%;  Score 12.6;  DB 6;  Length 57;
  Best Local Similarity   78.9%;  Pred. No. 2.5e-04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 ccgaatcaaatgttatga 20
           ||  ||||| || ||||
Db       4 CCTGGTCAAGCCTGGTGGA 24


RESULT  14
AX146941
LOCUS       AX146941        57 bp    DNA             PAT       08-JUN-2001
DEFINITION  Sequence 10 from Patent WO0134646.
ACCESSION   AX146941
VERSION     AX146941.1  GI:14446241
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 57)
  AUTHORS   Chang,P.C., Kivirikko,K.I., Neil,I.R., Olsen,D.R. and Polarek,J.W.
  TITLE     Recombinant relaxins
  JOURNAL   Patent: WO 0134646-A 10 17-MAY-2001;
            FIBROGEN, INC. (US)
FEATURES             Location/Qualifiers
     source          1..57
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        9 a      12 c      20 g      16 t
ORIGIN


  Query Match             63.0%;  Score 12.6;  DB 6;  Length 57;
  Best Local Similarity   78.9%;  Pred. No. 2.5e-04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 ccgaatcaaatgttatga 20
           ||  ||||| || ||||
Db       4 CCTGGTCAAGCCTGGTGGA 24


RESULT  15
MUSIGHEAA
LOCUS       MUSIGHEAA       60 bp    DNA             ROD       27-APR-1993
DEFINITION  Mouse Ig germline H chain gene, S-epsilon region.
ACCESSION   M20879
VERSION     M20879.1  GI:190150
KEYWORDS    germline; immunoglobulin heavy chain; immunoglobulin epsilon;
            switch region.
SOURCE      Mouse DNA.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Yoshida,K., Matsuoka,M., Usuda,S., Mori,A., Ishizaka,K. and
            Sakano,H.
  TITLE     Immunoglobulin switch circular DNA in the mouse infected with
            Nippostrongylus brasiliensis: Evidence for successive class
            switching from mu to epsilon via gamma 1
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 87, 7829-7833 (1990)
  MEDLINE   91045601
FEATURES             Location/Qualifiers
     source          1..60
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /sub_species="domesticus"
                     /db_xref="taxon:10090"
     misc_feature    1..59
                     /partial
                     /note="S-epsilon region"
BASE COUNT       11 a      9 c      22 g      18 t
ORIGIN


  Query Match             63.0%;  Score 12.6;  DB 10;  Length 60;
  Best Local Similarity   78.9%;  Pred. No. 2.5e-04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 ccgaatcaaatgttatga 20
           ||  ||||| || ||||
Db      16 CCGAGTCAAGCCTGGTGGA 34


Search completed: February  9, 2002, 06:24:01
Job time: 15781 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

SM nucleic - nucleic search, using sw model

Run on: February 9, 2002, 06:14:08 ; Search time 1863.18 Seconds
(without alignments)
177.086 Million cell updates/sec

Title: US-10-006-430-88
Perfect score: 20
Sequence: 1 gtgaccgtaaagtttcgtaac 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 147240 seqs, 82489825 residues

Total number of hits satisfying chosen parameters: 986436

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_os:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_os:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_mus:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_mus:*
35: em_htg_rod:*
36: em_htg_other:*

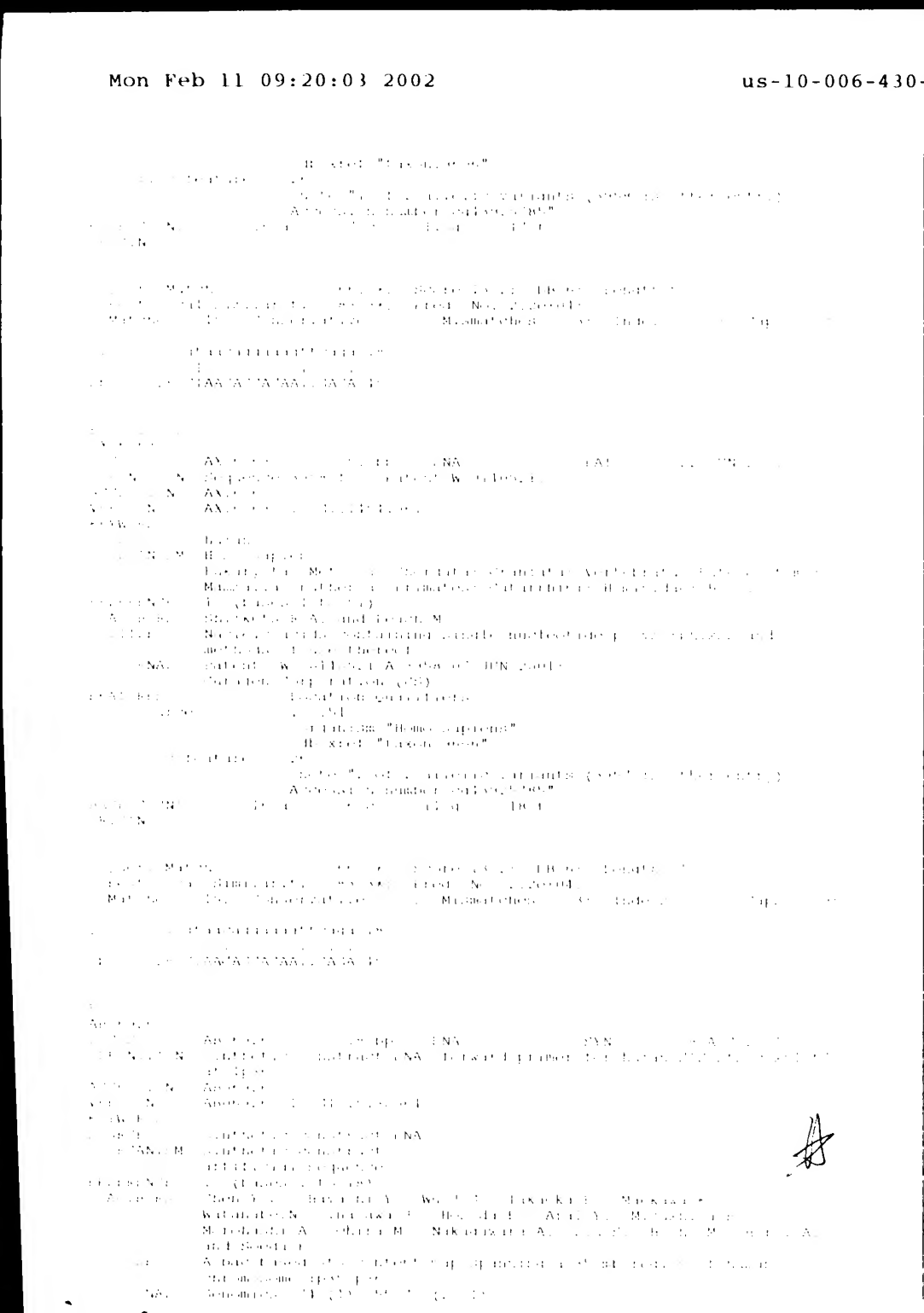
Hit No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

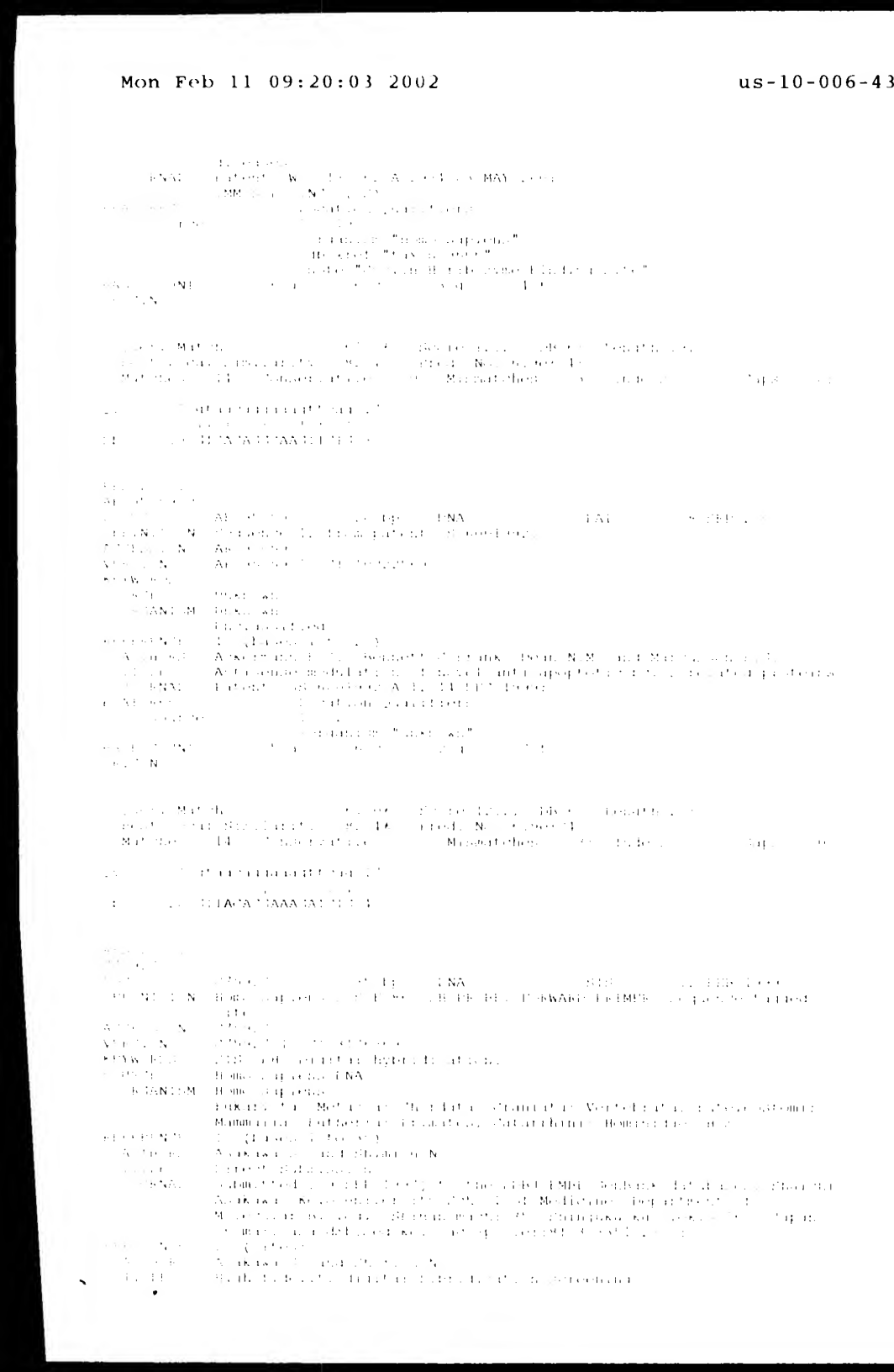
Result No.	Score	Query Match	Length	DB	ID	Description
1	13.2	66.0	51	6	AX160629	AX160629 Sequence
2	13.2	66.0	51	6	AX160630	AX160630 Sequence
3	13	65.0	18	12	AB069260	AB069260 Staphylo
4	12.8	64.0	19	6	AC0306	AC030[illegible] Sequence
c 5	12.6	63.0	45	6	AX127327	AX127327 Sequence
c 6	12.6	63.0	45	6	AX127328	AX127328 Sequence
c 7	12.4	62.0	22	4	DE620541402	178500 Vaccinia [illegible]
c 8	12.4	62.0	40	6	AX025232	AX025232 Sequence
c 9	12.2	61.0	19	6	AX131716	AX131716 Sequence
c 10	12.2	61.0	20	6	AR095059	AR095059 Sequence
11	12.2	61.0	35	11	C75927	C75927 Hemoglobin [illegible]
c 12	12.2	61.0	51	6	AX156944	AX156944 Sequence
13	12.2	61.0	59	9	BSA241953	A[illegible] Bovine [illegible]
14	12	60.0	28	6	143628	143628 Sequence 3
c 15	12	60.0	30	6	143627	143627 Sequence 2
16	12	60.0	36	6	AX061261	AX061261 Sequence
c 17	12	60.0	36	6	AX061262	AX061262 Sequence
c 18	12	60.0	50	6	AX159140	AX159140 Sequence
19	12	60.0	51	6	AX116001	AX116001 Sequence
20	12	60.0	51	6	AX157014	AX157014 Sequence
c 21	12	60.0	51	6	AX159137	AX159137 Sequence
c 22	12	60.0	51	6	AX159139	AX159139 Sequence
23	12	60.0	60	6	AR051681	AR051681 Sequence
24	11.8	59.0	17	6	E04837	E04837 Synthetic [illegible]
c 25	11.8	59.0	41	6	AR121249	AR121249 Sequence
26	11.8	59.0	36	6	AX140180	AX140180 Sequence
c 27	11.8	59.0	37	6	AR022681	AR022681 Sequence
c 28	11.8	59.0	37	6	AR037596	AR037596 Sequence
c 29	11.8	59.0	39	6	AR022568	AR022568 Sequence
c 30	11.8	59.0	39	6	AR037583	AR037583 Sequence
c 31	11.8	59.0	39	6	AR083834	AR083834 Sequence
c 32	11.8	59.0	40	6	AR022573	AR022573 Sequence
c 33	11.8	59.0	40	6	AR024250	AR024250 Sequence
c 34	11.8	59.0	40	6	AR026918	AR026918 Sequence
c 35	11.8	59.0	40	6	AR037588	AR037588 Sequence
c 36	11.8	59.0	40	6	AR045103	AR045103 Sequence
c 37	11.8	59.0	40	6	AR091804	AR091804 Sequence
c 38	11.8	59.0	40	6	E05494	E05494 DNA sequence
c 39	11.8	59.0	40	22	E11592	E11592 [illegible] primer
c 40	11.8	59.0	50	6	A38858	A38858 Sequence 12
c 41	11.8	59.0	50	6	A73047	A73047 Sequence 23
c 42	11.8	59.0	50	6	A73139	A73139 Sequence 23
c 43	11.8	59.0	50	6	AR013763	AR013763 Sequence
c 44	11.8	59.0	50	6	AR126943	AR126943 Sequence
45	11.8	59.0	51	6	AX161811	AX161811 Sequence

ALIGNMENTS

RESULT 1
AX160629
LOCUS AX160629 51 bp DNA PAT 21-JUN-2001
DEFINITION Sequence 957 from Patent WO0140521.
ACCESSION AX160629
VERSION AX160629.1 GI:14541960
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 51)
AUTHORS Shimkets,R.A. and Leach,M.
TITLE Nucleic acids containing simple nucleotide polymorphisms and methods of use thereof
JOURNAL Patent: WO 0140521-A 957 07-JUN-2001;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source 1..51
/organism="Homo sapiens"



```
  AUTHORS   Strosberg,A.D. and Eshdat,Y.
  TITLE     Plant enzymes with phospholipid hydroperoxide glutathione
            peroxidase activity, their analogues and their use
  JOURNAL   Patent: EP 1111055 A 40 27 JUN 2001;
            Verigen (FR)
FEATURES             Location/Qualifiers
     source          1..45
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="primers"
BASE COUNT        14 a      10 c       9 g      12 t
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 6;  Length 45;
  Best Local Similarity   78.9%;  Prob. No. 4.5e-04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       7 tacaaaaattcaaac 20
           ||||||| ||| | | |
Db      22 IAAC'AGGAC'AGIIAAC'AGC 4


RESULT   6
AX127428/c
LOCUS       AX127428        45 bp    DNA             PAT       15 MAY 2001
DEFINITION  Sequence 48 from Patent EP1111055.
ACCESSION   AX127428
VERSION     AX127428.1  GI:14143475
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Strosberg,A.D. and Eshdat,Y.
  TITLE     Plant enzymes with phospholipid hydroperoxide glutathione
            peroxidase activity, their analogues and their use
  JOURNAL   Patent: EP 1111055 A 40 27 JUN 2001;
            Verigen (FR)
FEATURES             Location/Qualifiers
     source          1..45
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="primers"
BASE COUNT        12 a      10 c       9 g      14 t
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 6;  Length 45;
  Best Local Similarity   78.9%;  Prob. No. 4.5e-04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       7 tacaaaaattcaaac 20
           ||||||| ||| | | |
Db      22 IAAC'AGGAC'AGIIAAC'AGC 4


RESULT   7
IX61716/c
LOCUS       IX61716         22 bp    DNA             MAM       29 NOV 1996
DEFINITION  Canis familiaris (    ) 20540) DNA, SES primer.
ACCESSION   I 06500
VERSION     I 06500.1  GI:1147[illegible]
KEYWORDS    genetic marker; m    satellite; tetranucleotide repeat.
SOURCE      dog
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Francisco,L.V., Langston,A.A., Mellersh,C.S., Neal,C.L. and
            Ostrander,E.A.
  TITLE     A class of highly polymorphic tetranucleotide repeats for canine
            genetic mapping
  JOURNAL   Mamm. Genome 7 (5), 359-362 (1996)
  MEDLINE   96269693
FEATURES             Location/Qualifiers
     source          1..22
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
                     /clone="20548"
     primer_bind     1..22
                     /note="20548"
                     /evidence=experimental
BASE COUNT         4 a       5 c       4 g       9 t
ORIGIN

  Query Match             62.0%;  Score 12.4;  DB 4;  Length 22;
  Best Local Similarity   92.9%;  Prob. No. 5.6e-04;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       7 qaaaaattcaaac 20
           ||||||||| |||
Db      21 GGAAAGTTCAAAC 8


RESULT   8
AX025232/c
LOCUS       AX025232        40 bp    DNA             PAT       [illegible]
DEFINITION  Sequence 15 from Patent DE19849348.
ACCESSION   AX025232
VERSION     AX025232.1  GI:10180946
KEYWORDS    .
SOURCE      rhesus monkey.
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Schmidt,M. and Kolle,C.V.
  JOURNAL   Patent: DE 19849348 A 15 27 APR 2000;
            UNIV LUDWIGS ALBERT (DE)
FEATURES             Location/Qualifiers
     source          1..40
                     /organism="Macaca mulatta"
                     /db_xref="taxon:9544"
     UTR             31..40
                     /note="Die ersten 10 Nucleotide von pEN [illegible]
                     (NCBI/Genbank: M28345)"
BASE COUNT        12 a       9 c       9 g      10 t
ORIGIN

  Query Match             62.0%;  Score 12.4;  DB 6;  Length 40;
  Best Local Similarity   92.9%;  Prob. No. 5.6e-04;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 taacaaaaattcaaac 15
           ||||||| ||||
Db      25 IAAC'AGGAC'AGTTC 12


RESULT   9
AX131716/c
LOCUS       AX131716        19 bp    DNA             PAT       15 MAY 2001
DEFINITION  Sequence 2954 from Patent WO0150462.
ACCESSION   AX131716
VERSION     AX131716.1  GI:14138021
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   Robbins,J.M. and Britz,R.
  TITLE     Ribozyme therapy for the treatment of proliferative [illegible] and eye
```



```
  AUTHOR(S)   Scalia,C.
  TITLE       Direct Submission
  JOURNAL     Submitted (27 APR 1995) Scalia C., Heim Pal Pediatric Hospital
              Budapest, Bokay u. 53-54, H-1958 Hungary
COMMENT       Related sequence M77430.
FEATURES             Location/Qualifiers
     source          1..30
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="Y"
                     /map="Yq11"
     misc_feature    1..30
                     /gene="spl40"
                     /note="spl40"
     misc_feature    5..29
                     /gene="spl40"
                     /number=4
BASE COUNT        16 a        5 c        7 g        2 t
ORIGIN


  Query Match             83.0%;  Score 124.5;  DB 9;  Length 30;
  Best Local Similarity   86.4%;  Pred. No. 7.3e-04;
  Matches   19;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        1 gtaacaggaaaattcaaaac 20
            ||| ||| ||| ||||| |||
Db        4 GTAACAGAAAAATTCAAC 23


RESULT  14
L36628
LOCUS       L36628                28 bp    DNA             PAT       06-FEB-1997
DEFINITION  Sequence 8 from patent US 5593840.
ACCESSION   L36628
VERSION     L36628.1  GI:1824419
KEYWORDS    .
SOURCE      Unknown
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 28)
  AUTHORS   Bhatnagar,S.K., George,A.L. Jr. and Nazarenko,I.
  TITLE     Amplification of nucleic acid sequences
  JOURNAL   Patent: US 5593840 A 8 14-JAN-1997;
FEATURES             Location/Qualifiers
     source          1..28
                     /organism="unknown"
BASE COUNT        12 a        7 c        3 g        6 t
ORIGIN


  Query Match             60.0%;  Score 12;  DB 6;  Length 28;
  Best Local Similarity   75.0%;  Pred. No. 9e+04;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        1 gtaacaggaaaattcaaaac 20
            | || ||  || ||| ||||
Db        6 GCAATAGAAAACTTCCGAAC 28


RESULT  15
L36627
LOCUS       L36627                40 bp    DNA             PAT       06-FEB-1997
DEFINITION  Sequence 7 from patent US 5593840.
ACCESSION   L36627
VERSION     L36627.1  GI:1824418
KEYWORDS    .
SOURCE      Unknown
  ORGANISM  Unknown.
            Unclassified.

REFERENCE   1  (bases 1 to 40)
  AUTHORS   Bhatnagar,S.K., George,A.L. Jr. and Nazarenko,I.
  TITLE     Amplification of nucleic acid sequences
  JOURNAL   Patent: US 5593840 A 7 14-JAN-1997;
FEATURES             Location/Qualifiers
     source          1..40
                     /organism="unknown"
BASE COUNT        9 a        5 c        8 g        12 t
ORIGIN


  Query Match             60.0%;  Score 12;  DB 6;  Length 40;
  Best Local Similarity   75.0%;  Pred. No. 9e+04;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        1 gtaacaggaaaattcaaaac 20
            | || ||  || ||| ||||
Db       20 GCAATAGAAAACTTCCGAAC 1


Search completed: February  9, 2002, 06:24:09
Job time: 16789 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 8, 2002, 18:56:25 ; Search time 8252.33 Seconds
(without alignments)
26,043 Million cell updates/sec

Title: US-10-006-430-49
Perfect score: 60
Sequence: 1 caccacagcaccatgctcag 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1134937 seqs, 537280281 residues

Total number of hits satisfying chosen parameters: 111874

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estmam:*
5: em_estpri:*
6: em_estrod:*
7: em_estvrt:*
8: em_estvrx:*
9: em_htc:*
10: qb_est1:*
11: qb_est2:*
12: qb_htc:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_ :*
17: em_gss_pro:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Prob. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% query Match	Length	DB	ID	Description
1	13.6	68.0	50	13	AU104165	AU104165 AU104165
2	13.4	67.0	33	13	AZ608228	AZ608228 IM0432111
3	13	65.0	19	13	AZ465883	AZ465883 IM0276807
4	12.6	63.0	41	11	BI9040	BI9040 yq21d12.r1
5	12.6	63.0	47	13	AZ494996	AZ494996 IM0330C07
6	12.6	63.0	50	10	AU106824	AU106824 AU106824
7	12.2	61.0	28	13	AZ843553	AZ843553 IM0317N13
8	12.2	61.0	38	13	AZ594430	AZ594430 IM0403B22
9	12.2	61.0	43	11	W62348	W62348 wp99e05.r1
10	12.2	61.0	43	13	AZ375966	AZ375966 IM0129C07
11	12.2	61.0	49	10	AA641790	AA641790 np68102.s
12	12.2	61.0	60	13	Q3A50611	AJ242117 Gallus ga
13	12	60.0	25	10	AI155015	AI155015 AI155015
14	12	60.0	33	13	AZ819863	AZ819863 IM0543B24
15	12	60.0	46	10	AA958824	AA958824 ww52d12.s
16	12	60.0	49	10	AA743162	AA743162 ty17b05.s
17	12	60.0	50	10	AU104163	AU104163 AU104163
18	12	60.0	50	10	AU104164	AU104164 AU104164
19	12	60.0	55	11	BF527748	BF527748 602e09e5
20	12	60.0	55	13	AZ514470	AZ514470 IM0618B24
21	12	60.0	56	13	AZ583210	AZ583210 IM0728D10
22	12	60.0	56	13	B47105	B47105 RS 1-27 AI
23	12	60.0	60	10	AA758406	AA758406 ab11e11.s
24	11.8	59.0	26	13	FA129A07P	AI463677 yj bmocc
25	11.8	59.0	34	13	FA250P09Q	AI482061 yj bmocc
26	11.8	59.0	46	10	AA195618	AA195618 yz57c11
27	11.8	59.0	46	11	I56929	I56929 ya54d08.r2
28	11.8	59.0	47	13	AZ820416	AZ820416 IM0541D10
29	11.8	59.0	48	10	AA651618	AA651618 ns55a11.s
30	11.8	59.0	51	13	AZ489275	AZ489275 IM0521B04
31	11.8	59.0	55	10	AI147368	AI147368 ok96e10
32	11.8	59.0	55	10	AI744611	AI744611 wa14e09
33	11.8	59.0	60	10	AA816131	AA816131 xp44b12
34	11.6	58.0	28	10	AA027602	AA027602 no11d08.r1
35	11.6	58.0	33	13	AZ834747	AZ834747 IM0615C09
36	11.6	58.0	36	13	AZ625783	AZ625783 IM0477E11
37	11.6	58.0	42	13	AZ794999	AZ794999 IM0746D02
38	11.6	58.0	46	10	AI153694	AI153694 xz43d08.r1
39	11.6	58.0	46	13	TAB9D05Q	AI458004 1 brocc1
40	11.6	58.0	48	11	BG705927	BG705927 602e02d08
41	11.6	58.0	48	11	BG716497	BG716497 602e76a22
42	11.6	58.0	49	10	AA673527	AA673527 ns61c11
43	11.6	58.0	50	10	AU106809	AU106809 AU106809
44	11.6	58.0	50	10	AU106810	AU106810 AU106810
45	11.6	58.0	50	10	AU106811	AU106811 AU106811

ALIGNMENTS

RESULT 1
AU104165/c
LOCUS AU104165 50 bp mRNA EST 08-APR-2001
DEFINITION AU104165 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone HEP15569, mRNA sequence.
ACCESSION AU104165
VERSION AU104165.1 GI:13553686
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Tsunoda,T., Saito,R., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Suyama,M., Okubo,K., Sakaki,Y. and Sugano,S.
TITLE Fine Structural analysis of transcription start sites in human mRNAs using full length enriched and 5' end enriched cDNA libraries
JOURNAL Unpublished (2001)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length enriched and a 5' end enriched cDNA library. Gene 200 (1-2), 149-156 (1997)
FEATURES Location/Qualifiers
source 1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP15569"
/clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT 11 a 16 c 18 g 5 t
ORIGIN

```
BASE COUNT        4 a      6 c      3 g      4 t
ORIGIN

  Query Match             65.0%;  Score 13;  DB 13;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 5.1e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 acagcacacatgct 17
           ||||||||||||||
Db       7 ACAGCACCATGCT 19


RESULT   3
BI004o
LOCUS       BI004o          41 bp    mRNA            EST       14 APR 1999
DEFINITION  yz26d12.r1 Soares infant brain 1NIB Homo sapiens cDNA clone
            IMAGE:3 265 5', similar to tr:X65664 SOX 8 PROTEIN (HUMAN);; mRNA
            sequence.
ACCESSION   BI004o
VERSION     BI004o.1  GI:7726o6
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 41)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Lee,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The Washu Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 2689
            High quality sequence start: 1 High quality sequence stop: 1
            Source: IMAGE Consortium, LLNL This clone is available royalty free
            through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov)
            for further information. Trace considered overall poor quality
            Insert Length: 2689  Std Error: 0.00
            Seq primer: M13REV
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1..41
                     /organism="Homo sapiens"
                     /db_xref="GDB:4056612"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3265"
                     /clone_lib="Soares infant brain 1NIB"
                     /sex="female"
                     /dev_stage="2 days post natal"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: whole brain; Vector: Lafmid BA; Site_1: Not
                     I; Site_2: Xho I/Sal I; 1st strand cDNA was primed with a Not
                     I   oligo(dT) primer [5'
                     AACTGGAACAAIICTCCCCCCCCAGGAATTTTTTTTTTTTTTTT 3'];
                     double stranded cDNA was ligated to Xho I/Sal I adaptors
                     (Pharmacia), digested with Not I and directionally cloned
                     into the Not I and Xho I/Sal I sites of the Lafmid BA vector.
                     Library went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       11 a     15 c     10 g      4 t       1 others
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 11;  Length 41;
  Best Local Similarity   75.0%;  Pred. No. 8.7e+04;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 cagcacacacatgctcaag 20
           | |||| | ||  ||||||
Db       7 CTGCACAACGCCNAGCTCAG 26


RESULT   5
AZ494996
LOCUS       AZ494996        47 bp    DNA             GSS       17-OCT-20
DEFINITION  IM0330707R Mouse 10kb plasmid 000C1M library Mus musculus genomic
            clone 0000IM0330707 R, DNA sequence.
ACCESSION   AZ494996
VERSION     AZ494996.1  GI:10670111
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Dunn,D., Aoyagi,A., Barbet,M., Beaucern,L., Duval,B., Bass,T.,
            Islam,H., Iamarre,S., Mahmood,M., Mermon,E., Pederson,J., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert R. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 508, Biomedical Polymers Research Bldg., 20 S. 2030 E.  Slc, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddnn@genetics.utah.edu
            Insert Length: 10000  Std Error: 0.00
            Plate: 0330  row: C  column: 07
            Seq primer: CACACAGGAAACAGCTATGAC
            Class: plasmid ends
            High quality sequence stop: 47.
FEATURES             Location/Qualifiers
     source          1..47
                     /organism="Mus musculus"
                     /strain="c57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="0000IM0330707"
                     /clone_lib="Mouse 10kb plasmid 000C1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10 Gold, ti resistant', 1 "
                     /note="Vector: PWM2nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high molar excess. The
                     adaptored DNA was purified and size selected 1 to 6.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWM2 (gi|4732114|gb|AF129072.1), a copy number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically competent E. coli XL10 Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       17 a     11 c      9 g     10 t
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 13;  Length 47;
  Best Local Similarity   78.9%;  Pred. No. 8.9e+04;
```

```
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy [illegible]
Db [illegible] AAAAAA AAA [illegible]

RESULT [illegible]
[illegible]
LOCUS       [illegible] mRNA [illegible]
DEFINITION  [illegible] Homo sapiens cDNA library [illegible] cDNA clone
            KAT[illegible] mRNA sequence.
ACCESSION   [illegible]
VERSION     [illegible]
KEYWORDS    [illegible]
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to [illegible])
  AUTHORS   Suzuki,Y., [illegible], Hata,
            [illegible], Nakamura,Y., Morishita,S., [illegible]
            K., Suyama,A. and Sugano,S.
  TITLE     Fine Structural analysis of transcription start sites of human
            mRNA [illegible] full length enriched and 5' end enriched cDNA libraries
  JOURNAL   Unpublished [illegible]
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            [illegible] Minato-ku, Tokyo [illegible] Japan
            [illegible]
            Suzuki,Y., Yoshitomo, Nakagawa,K., Maruyama,K., [illegible] and Sugano
            [illegible] Construction and characterization of a full length enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1..[illegible]
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="KAT[illegible]"
                     /note="Homo sapiens cDNA [illegible]"
BASE COUNT  [illegible]
ORIGIN

  Query Match             [illegible]  Score [illegible]  DB [illegible]  Length [illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches  [illegible]  Conservative  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy [illegible]
Db [illegible] TATATAAAT AATAAA [illegible]

RESULT [illegible]
[illegible]
LOCUS       [illegible] DNA [illegible] GSS [illegible]
DEFINITION  [illegible] Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone [illegible] DNA sequence.
ACCESSION   [illegible]
VERSION     [illegible]
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to [illegible])
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beaudoin,L., Duval,B., [illegible]
            Islam,H., [illegible], Mahmoud,M., Meenen,E., Pedersen,[illegible]
            [illegible], Stokes,R., Tingey,A., [illegible]
            and Wright,D., Weiss,R.
  TITLE     Mouse whole genome [illegible] with paired end reads [illegible] 10kb
            plasmid inserts
  JOURNAL   Unpublished [illegible]
```

```
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            [illegible] Biomedical Polymers Research Bldg [illegible] E., SLC, UT
            84112, USA
            Tel: [illegible]
            Fax: [illegible]
            Email: [illegible]
            Insert length: 10000 [illegible]
            Plate: [illegible] row: N column: [illegible]
            Seq primer: [illegible]AAAA[illegible]
            Class: plasmid ends
            High quality sequence stop: [illegible]
FEATURES             Location/Qualifiers
     source          1..[illegible]
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M[illegible]"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. coli strain XL10-Gold [illegible]"
                     /note="Vector: pWD42[illegible]; Purified genomic [illegible] from M.
                     musculus C57BL/6J (male) was obtained from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/[illegible]). The DNA
                     was hydrodynamically sheared by repeated [illegible] through a
                     [illegible] orifice at constant velocity. [illegible] sheared DNA
                     was blunt end repaired with T4 DNA [illegible] and T4
                     polynucleotide kinase. Adaptor oligo [illegible] were
                     ligated to the blunt ends in high [illegible]. The
                     adaptored DNA was purified and size selected [illegible]
                     [illegible] kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared [illegible] a derivative
                     of pWD42 (gi|[illegible]|gb|AF12[illegible]) [illegible]
                     [illegible] derivative of plasmid [illegible]. The [illegible] was ligated
                     with adaptors complementary to the insert [illegible] and
                     purified. The sheared, adaptored mouse DNA [illegible] annealed to
                     adaptored vector DNA, and transformed into
                     chemically competent E. coli XL10-Gold ([illegible]) cells
                     and selected for ampicillin resistance."
BASE COUNT  [illegible]
ORIGIN

  Query Match             [illegible]  Score [illegible]  DB [illegible]  Length 28
  Best Local Similarity   82.4%  Pred. No. [illegible]
  Matches   14;  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy [illegible]
Db [illegible] GCAGAATATATAT[illegible]GA 25
```

```
  JOURNAL   Unpublished (2000)
  COMMENT   Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 00N, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 3539
            Email: rbw@genetics.utah.edu
            Insert length: 10000  Std Error: 0.00
            Plate: 0103  row: B  column: 22
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop:  98.
FEATURES             Location/Qualifiers
     source          1..38
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UI0C1M0403B22"
                     /clone_lib="Mouse 10kb plasmid UI0C1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10 Gold, ri resistant, F'"
                     /note="Vector: PWB42B2; Purified genomic DNA from M
                     musculus (C57BL/6J male) was obtained from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/imr/).  The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWB42 (gi|4741211|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically competent E. coli XL10 Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT         8 a        7 c       14 g       10 t
ORIGIN


  Query Match             61.0%;  Score 12.2;  DB 13;  Length 38;
  Best Local Similarity   82.4%;  Prob. No. 1.2e+05;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 cgcaacgcaccatgct 17
           ||||||| ||  ||||
Db      25 CGACCACGGACTTGCT 10


RESULT  9
W62458
LOCUS       W62458             44 bp    mRNA            EST       07-JUN-1996
DEFINITION  md96e05.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone IMAGE:386120 5' similar to SW:KGUA_PIG P41006 GUANYLATE
            KINASE ;, mRNA sequence.
ACCESSION   W62458
VERSION     W62458.1  GI:1369153
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,P.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Stepto e,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            Washu-HHMI Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:2379952
            Trace considered overall poor quality
            Possible reversed clone; similarity on wrong strand
            Seq primer: mob.REGA+EF
            High quality sequence stop:  1.
FEATURES             Location/Qualifiers
     source          1..43
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:386120"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5 14.5dpc total fetus"
                     /lab_host="DH10B"
                     /note="Vector: pT713 Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; site 2: Eco RI; 1st strand cDNA
                     was primed with a Not I  oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGGAAATTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 x 13dpc and 2 x
                     14.5dpc embryos (total RNA provided by  Mineri K., Wayne
                     State Univ., from 2 l; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT713 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo. "
BASE COUNT         11 a       14 c        7 g       11 t
ORIGIN


  Query Match             61.0%;  Score 12.2;  DB 11;  Length 43;
  Best Local Similarity   82.4%;  Prob. No. 1.3e+05;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 agcacgcaccatgctc 18
           |||||||| ||  ||
Db       6 AGCACAGCAGCATCTC 22


RESULT  10
AZ375966
LOCUS       AZ375966           43 bp    DNA             GSS       12-OCT-2001
DEFINITION  IM0129C07R Mouse 10kb plasmid UI0C1M library Mus musculus genomic
            clone UI0C1M0129C07 R, DNA sequence.
ACCESSION   AZ375966
VERSION     AZ375966.1  GI:10489666
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 43)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Bescorn,L., Brzal,P., Bamii,C.,
            Islam,H., Jomaacre,S., Mahmood,M., Meetoen,E., Pedersen,L., Reilly
            ,M., Rose,M., Rose,R., Stokes,K., Tingey,A., von Niederhausern,A.
            and Wright,D., Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
```

```
     source          1. .60
                     /organism "Gallus gallus"
                     /db_xref "taxon:9031"
                     /chromosome "Z"
                     /map "Zq"
BASE COUNT       11 a      19 c      18 g      12 t
ORIGIN


  Query Match             71.0%;  Score 12.2;  DB 13;  Length 60;
  Best Local Similarity   82.4%;  Pred. No. 1.4e+05;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 gcccagcaccatgctcca 19
           ||||||||||||||||||
Db      17 GCCCAGCACCATGCTCCA 34


RESULT  13
AZ819863/c
LOCUS       AZ819863          34 bp    DNA             GSS       20-FEB-2002
DEFINITION  2M009d0148 Mouse 10kb plasmid 009CIM library Mus musculus genomic
            clone 009C2M009d0148 R, DNA sequence.
ACCESSION   AZ819863
VERSION     AZ819863.1  GI:12989771
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 34)
  AUTHORS   Dunn,D., Aoyari,A., Barber,M., Beacorn,I., Duval,I., Farr,T.,
            Islam,M., Longacre,S., Mahamoud,M., Meenan,E., Pedersen,J., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kt
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., Slc, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert length: 10000  Std Error: 0.00
            Plate: 0091  row: b  column: 14
            Seq primer: CACACAGGAAACAGCTATGAC
            Class: plasmid ends
            High quality sequence stop: 34.
FEATURES             location/qualifiers
     source          1. .34
                     /organism "Mus musculus"
                     /strain "C57BL/6J"
                     /db_xref "taxon:10090"
                     /clone "IMAGE:1476908"
                     /clone_lib "Soares mammary gland NMLMG"
                     /sex "female (lactating)"
                     /tissue_type "mammary gland"
                     /lab_host "DH10B"
                     /note "Vector: pT7T3 Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from mammary
                     gland tissue from a lactating female, and was then primed
                     with a Not I   oligo(dT) primer. Double stranded  cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I  and cloned into the Not I and Eco RI sites of the
                     modified pT3T3 vector.  Library is normalized.  Library
                     was constructed by Benito Soares and M. Fatima Bonaldo. "
BASE COUNT       9 a       6 c       12 g       7 t
ORIGIN


  Query Match             60.0%;  Score 12;  DB 13;  Length 25;
  Best Local Similarity   75.0%;  Pred. No. 1.4e+05;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 cagcacagcaccatgctccag 20
           | || ||||||||||  || |
Db      21 CGGCCCAGCACCATGCTCCGG 2


RESULT  14
AZ819864/c
LOCUS       AZ819864          33 bp    DNA             GSS       20-FEB-2002
DEFINITION  2M009d0148 Mouse 10kb plasmid 009CIM library Mus musculus genomic
            clone 009C2M009d0148 R, DNA sequence.
ACCESSION   AZ819864
VERSION     AZ819864.1  GI:12989771
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Dunn,D., Aoyari,A., Barber,M., Beacorn,I., Duval,I., Farr,T.,
            Islam,M., Longacre,S., Mahamoud,M., Meenan,E., Pedersen,J., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kt
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., Slc, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert length: 10000  Std Error: 0.00
            Plate: 0091  row: b  column: 14
            Seq primer: CACACAGGAAACAGCTATGAC
            Class: plasmid ends
            High quality sequence stop: 33.
FEATURES             location/qualifiers
     source          1. .33
                     /organism "Mus musculus"
                     /strain "C57BL/6J"
                     /db_xref "taxon:10090"
                     /clone "IMAGE:1476908"
                     /clone_lib "Soares mammary gland NMLMG"
                     /sex "female (lactating)"
                     /tissue_type "mammary gland"
                     /lab_host "DH10B"
                     /note "Vector: pT7T3 Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from mammary
                     gland tissue from a lactating female, and was then primed
                     with a Not I   oligo(dT) primer. Double stranded  cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I  and cloned into the Not I and Eco RI sites of the
                     modified pT3T3 vector.  Library is normalized.  Library
                     was constructed by Benito Soares and M. Fatima Bonaldo. "
BASE COUNT       9 a       4 c       11 g       9 t
ORIGIN
```

GenCore version 4.5
Copyright (c) 1[illegible] 2000 Compugen Ltd.

Run mode: nucleic search, using sw model

Run on: February 8, 2002, 19:38:52 ; Search time 2544.26 Seconds
(without alignments)
129.743 Million cell updates/sec

Title: US-10-006-430-30
Perfect score: 20
Sequence: 1 ccatcctgcaatatcatta 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 147240 seqs, 82485875 residues

Total number of hits satisfying chosen parameters: 586446

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_oz:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

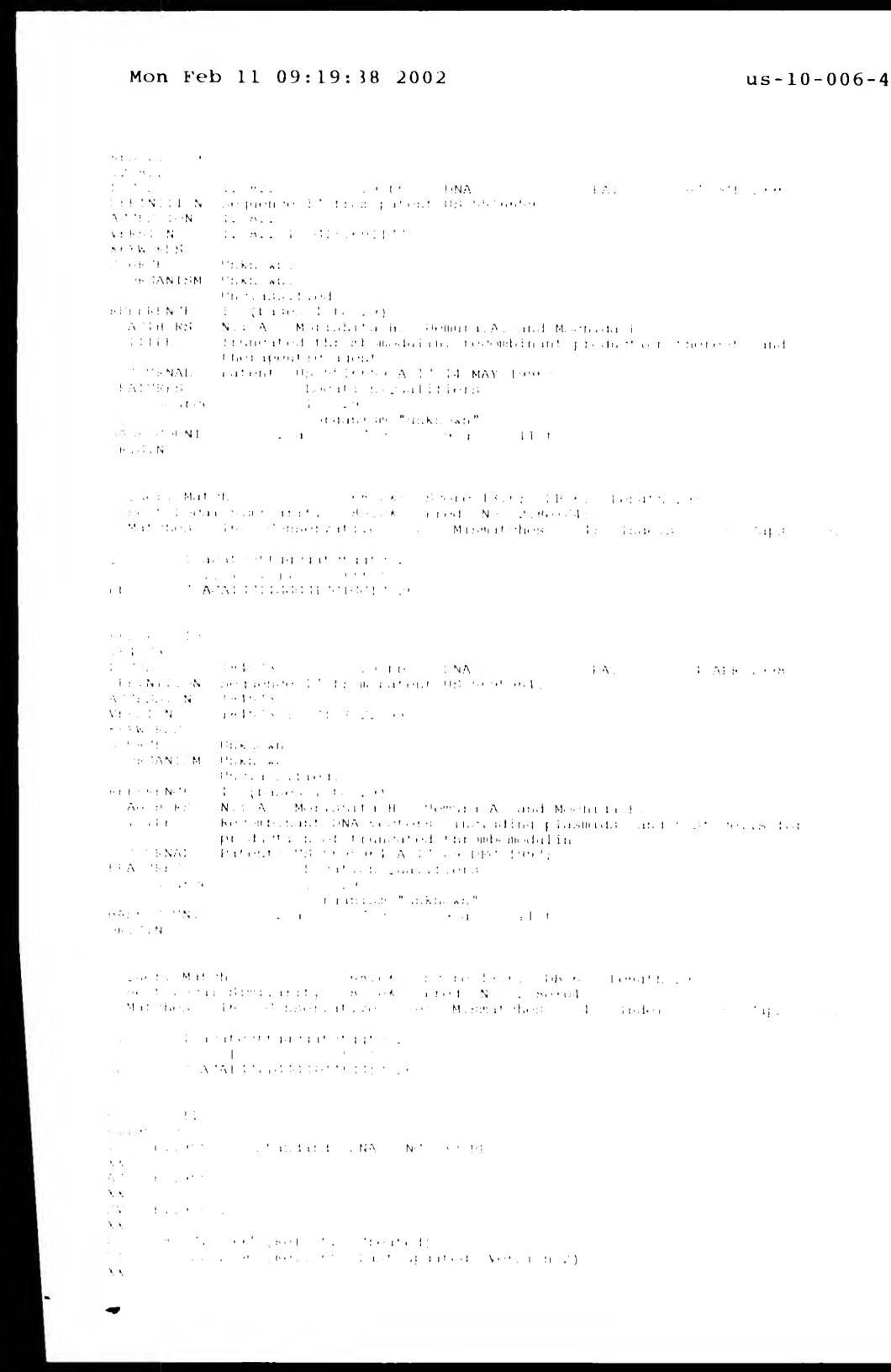
SUMMARIES

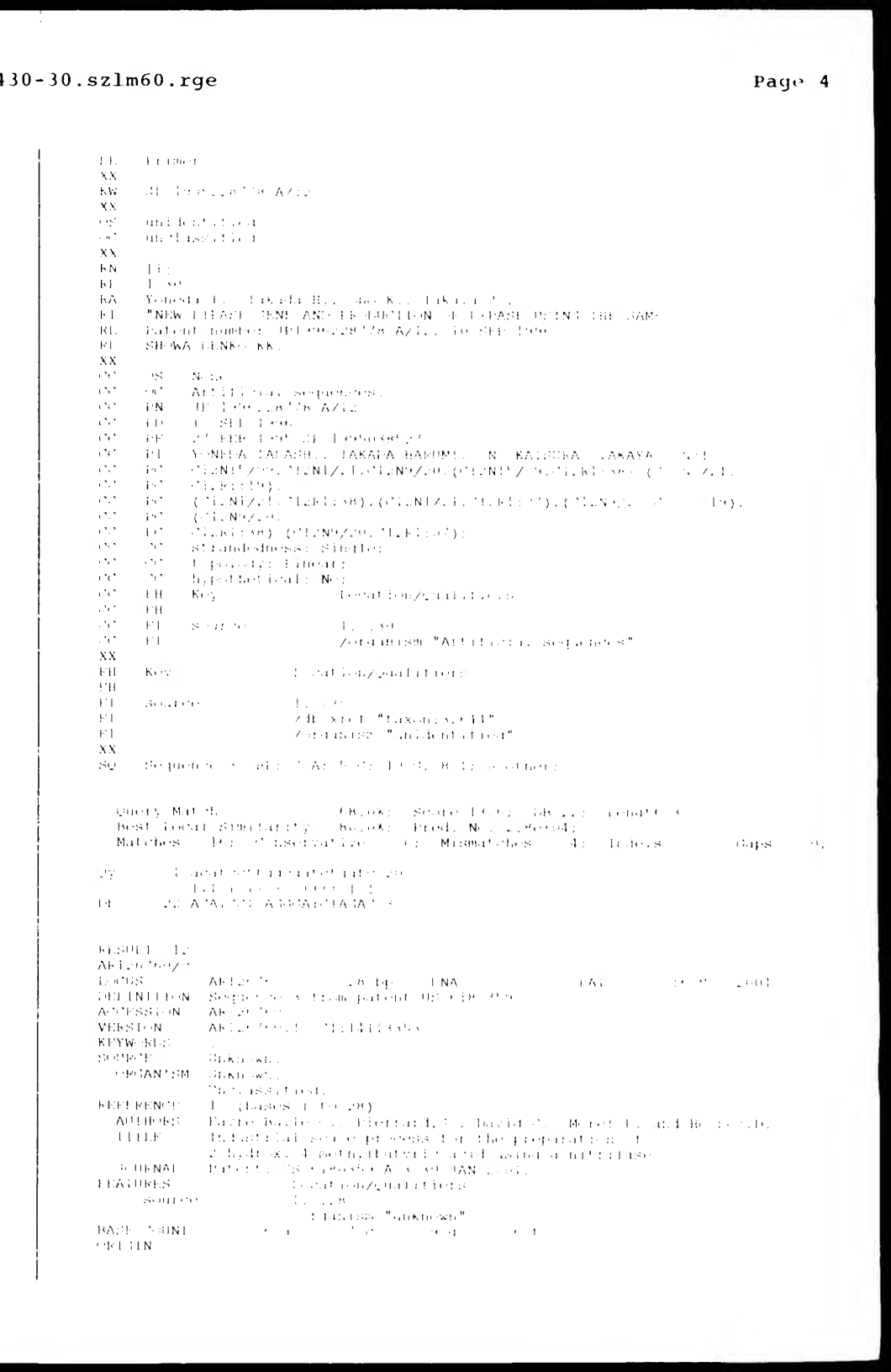
Result No.	Score	Query Match	Length	DB	ID	Description
c 1	15	75.0	41	6	AX004402	AX004402 Sequence 5
c 2	14.2	71.0	51	6	AX158801	AX158801 Sequence
c 3	14.2	71.0	51	6	AX158802	AX158802 Sequence
c 4	14.2	71.0	51	6	AX158803	AX158803 Sequence
c 5	14.2	71.0	51	6	AX158805	AX158805 Sequence
c 6	14.2	71.0	51	6	AX158806	AX158806 Sequence
7	14.2	71.0	53	6	E01772	E01772 DNA Sequence
c 8	14.2	71.0	54	6	E01773	E01773 DNA Sequence
9	13.6	68.0	29	6	120822	120822 Sequence 17
10	13.6	68.0	29	6	184573	184573 Sequence 17
c 11	13.6	68.0	30	22	E11955	E11955 Trimer, 9/2
c 12	13.4	67.0	28	6	AR126760	AR126760 Sequence
c 13	13.2	66.0	27	6	AR6957	AR6957 Sequence 48
c 14	13.2	66.0	27	6	AR048369	AR048369 Sequence
c 15	13.2	66.0	28	6	AX067982	AX067982 Sequence
c 16	13.2	66.0	51	6	AX165471	AX165471 Sequence
17	13	65.0	19	6	A64608	A64608 Sequence 27
18	12.6	63.0	22	6	A47958	A47958 Sequence 21
19	12.6	63.0	22	6	121763	121763 Sequence 21
c 20	12.6	63.0	25	6	AR081778	AR081778 Sequence
c 21	12.6	63.0	37	6	AX167839	AX167839 Sequence
22	12.6	63.0	49	6	AX167840	AX167840 Sequence
23	12.6	63.0	50	6	A19044	A19044 Nucleotide
c 24	12.6	63.0	50	6	A19045	A19045 Nucleotide
25	12.6	63.0	50	6	AX080812	AX080812 Sequence
c 26	12.6	63.0	51	6	AX158804	AX158804 Sequence
27	12.6	63.0	51	6	AX162515	AX162515 Sequence
28	12.6	63.0	51	6	AX162516	AX162516 Sequence
c 29	12.6	63.0	57	6	AR014089	AR014089 Sequence
c 30	12.6	63.0	57	6	106478	106478 Sequence 15
31	12.4	62.0	41	6	A18467	A18467 Tight chain
32	12.4	62.0	41	6	A24273	A24273 Interferon
33	12.4	62.0	41	6	AR028581	AR028581 Sequence
34	12.4	62.0	41	6	AR085792	AR085792 Sequence
35	12.2	61.0	24	6	AX079940	AX079940 Sequence
c 36	12.2	61.0	41	6	146795	146795 Sequence 5
37	12.2	61.0	41	6	146796	146796 Sequence 6
38	12.2	61.0	59	6	AR116174	AR116174 Sequence
39	12.2	61.0	59	6	AR126451	AR126451 Sequence
40	12.2	61.0	41	5	G08852	X05591 Clerkia pn
41	12.2	61.0	42	6	A47004	A47004 Sequence 6
42	12.2	61.0	42	6	A51242	A51242 Sequence 6
43	12.2	61.0	42	6	AR055031	AR055031 Sequence
44	12.2	61.0	42	6	AR108914	AR108914 Sequence
45	12.2	61.0	42	6	AR118542	AR118542 Sequence

ALIGNMENTS

RESULT 1
AX004402/c
LOCUS AX004402 41 bp DNA PAT 14-APR-2000
DEFINITION Sequence 5 from Patent WO9918198.
ACCESSION AX004402
VERSION AX004402.1 GI:9927869
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 41)
AUTHORS Grandi,G. and Abrignani,S.
TITLE Hepatitis c receptor protein cd81
JOURNAL Patent: WO 9918198-A 5 15-APR-1999;
CHIRON S.P.A. (IT); GRANDI GUIDO (IT)
FEATURES Location/Qualifiers
source 1..41
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligodeoxynucleotides"

```
                     /note "1 of 2 allelic variants (1644 is other entry)
                     Accession number cq42152942"
BASE COUNT        11 a      15 c      14 g      12 t
ORIGIN


  Query Match             71.0%;  Score 14.2;  DB 6;  Length 54;
  Best Local Similarity   84.2%;  Prod. No. 1.4e-04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 acatcettqqcatctqqt 19
           |||| |||| |||| |||
Db      21 ACATCCTTGCGCATCACGT 3


RESULT   6
AX158406/c
LOCUS       AX158406         51 bp    DNA             PAT       22-JUN-2001
DEFINITION  Sequence 1644 from Patent WO0140521.
ACCESSION   AX158406
VERSION     AX158406.1  GI:14539637
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Lench,N.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521 A 16-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..51
                     /db_xref="taxon:9606"
                     /organism="Homo sapiens"
     misc_feature    26
                     /note "2 of 2 allelic variants (1633 is other entry)
                     Accession number cq42152942"
BASE COUNT        11 a      16 c      13 g      11 t
ORIGIN


  Query Match             71.0%;  Score 14.2;  DB 6;  Length 51;
  Best Local Similarity   84.2%;  Prod. No. 1.4e-04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 acatcettqqcatctqqt 19
           |||| |||| |||| |||
Db      21 ACATCCTTGCGCATCACGT 3


RESULT   7
E01772/c
LOCUS       E01772           54 bp    DNA             PAT       29-SEP-1997
DEFINITION  DNA sequence of multi cloning site.
ACCESSION   E01772
VERSION     E01772.1  GI:2170025
KEYWORDS    JP 10882947914/A/1.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Kanda,F., Saigo,K. and Shibuya,I.
  TITLE     VECTOR, DETECTION OF GENE AND MANIFESTATION OF PROTEIN
  JOURNAL   Patent: JP 10882947914 A 1 01-DEC-1988;
            NIKKA OTSUKIT KK
COMMENT     OS   Artificial gene
            OC   Artificial sequence; Genes.
            PN   JP 10882947914/A
            PD   01-DEC-1988
            PF   27-MAY-1987 JP 1987130800
            PI   KANDA TOMOMASA, SAIGO KAORU, SHIBUYA ICHIRO
            PC   C12N15/00,C12P21/00,C12Q1/68,(C12P21/00,C12R1:19);
            Strandedness: double;
            CC   topology: linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key             Location/Qualifiers
            FH
            FT   misc_feature    1..54
            FT                   /note "multi cloning site"
FEATURES             Location/Qualifiers
     source          1..54
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        14 a      13 c      13 g      14 t
ORIGIN


  Query Match             71.0%;  Score 14.2;  DB 6;  Length 54;
  Best Local Similarity   84.2%;  Prod. No. 1.4e-04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 acatcettqqcatctqqt 19
           |||| |||| |||| |||
Db      13 ATATCCATGCGCATCTGCT 31


RESULT   8
E01773/c
LOCUS       E01773           54 bp    DNA             PAT       29-SEP-1997
DEFINITION  DNA sequence of multi cloning site.
ACCESSION   E01773
VERSION     E01773.1  GI:2170026
KEYWORDS    JP 10882947914/A/2.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Kanda,F., Saigo,K. and Shibuya,I.
  TITLE     VECTOR, DETECTION OF GENE AND MANIFESTATION OF PROTEIN
  JOURNAL   Patent: JP 10882947914-A 2 01-DEC-1988;
            NIKKA OTSUKIT KK
COMMENT     OS   Artificial gene
            OC   Artificial sequence; Genes.
            PN   JP 10882947914/A/2
            PD   01-DEC-1988
            PF   27-MAY-1987 JP 1987130800
            PI   KANDA TOMOMASA, SAIGO KAORU, SHIBUYA ICHIRO
            PC   C12N15/00,C12P21/00,C12Q1/68,(C12P21/00,C12R1:19);
            Strandedness: double;
            CC   topology: linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key             Location/Qualifiers
            FH
            FT   misc_feature    1..54
            FT                   /note "multi cloning site"
FEATURES             Location/Qualifiers
     source          1..54
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        12 a      14 c      14 g      14 t
ORIGIN


  Query Match             71.0%;  Score 14.2;  DB 6;  Length 54;
  Best Local Similarity   84.2%;  Prod. No. 1.4e-04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 acatcettqqcatctqqt 19
           |||| |||| |||| |||
Db      46 ATATCCATGCGCATCTGCT 28
```





```
  Query Match             67.0%;  Score 14.4;  DB 6;  Length 28;
  Best Local Similarity   93.3%;  Pred. No. 3.6e-04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       9 cttcgcgatcgtcgc 23
           |||| ||||||||||
Db      22 cttcACGAtCtGGtC 8


RESULT  14
AX067957/c
LOCUS       AX067957               27 bp    DNA             PAT       22-JAN-2000
DEFINITION  Sequence 48 from Patent WO9838306.
ACCESSION   AX067957
VERSION     AX067957.1  GI:6235741
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Bolbaum,G.
  TITLE     TRANSCRIPTS ENCODING IMMUNOREGULATORY POLYPEPTIDES
  JOURNAL   Patent: WO 9838306 A 48 03-SEP-1998;
            GENELABS TECH INC (US)
FEATURES             Location/Qualifiers
     source          1..27
                     /organism="unidentified"
                     /isolate="PRIMER CA4 10"
                     /db_xref="taxon:32644"
BASE COUNT         6 a        8 c        8 g        5 t
ORIGIN


  Query Match             66.0%;  Score 14.2;  DB 6;  Length 27;
  Best Local Similarity   83.3%;  Pred. No. 4.5e-04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 catcgttgcgatcgat 17
           |||| ||| || ||||
Db      18 CATCGGTGCGCAACTGGT 3


RESULT  14
AR048369/c
LOCUS       AR048369               27 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 48 from patent US 5821091.
ACCESSION   AR048369
VERSION     AR048369.1  GI:5920712
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Bolbaum,G.
  TITLE     Method of identifying activated T cells
  JOURNAL   Patent: US 5821091 A 48 13-OCT-1998;
FEATURES             Location/Qualifiers
     source          1..27
                     /organism="unknown"
BASE COUNT         6 a        8 c        8 g        5 t
ORIGIN


  Query Match             66.0%;  Score 14.2;  DB 6;  Length 27;
  Best Local Similarity   83.3%;  Pred. No. 4.5e-04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 catcgttgcgatcgat 17
           |||| ||| || ||||
Db      18 CATCGGTGCGCAACTGGT 3


RESULT  15
```

```
AX067982/c
LOCUS       AX067982               28 bp    DNA             PAT       22-JAN-2000
DEFINITION  Sequence 27 from Patent WO0077185.
ACCESSION   AX067982
VERSION     AX067982.1  GI:12329796
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 28)
  AUTHORS   Levin,J.Z., Banet,M.W. and Zhang,F.
  TITLE     Herbicide target genes and methods
  JOURNAL   Patent: WO 0077185 A 27 21-DEC-2000;
            Novartis AG (CH) ; Novartis Erfindungen Verwaltungsgesellschaft
            m.b.H. (AT)
FEATURES             Location/Qualifiers
     source          1..28
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="oligonucleotide"
BASE COUNT         7 a        9 c        5 g        7 t
ORIGIN


  Query Match             66.0%;  Score 14.2;  DB 6;  Length 28;
  Best Local Similarity   83.3%;  Pred. No. 4.5e-04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 acatcgttgcgatcgat 18
           | |||| ||| || ||||
Db      19 ATATCCATGCGCAACTGGT 2


Search completed: February  8, 2002, 19:48:53
Job time: 13780 sec
```